(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
18 April 2002 (18.04.2002)

PCT

(10) International Publication Number
WO 02/31511 A2

(51) International Patent Classification[7]: G01N 33/68

(21) International Application Number: PCT/US01/32047

(22) International Filing Date: 12 October 2001 (12.10.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/239,750 12 October 2000 (12.10.2000) US

(71) Applicant: ICOS CORPORATION [US/US]; 22021 20th Avenue, S.E., Bothell, WA 98021 (US).

(72) Inventor: STAUTON, Donald, E.; 6502 113th Avenue NE, Kirkland, WA 98033 (US).

(74) Agent: WILLIAMS, Joseph, A., Jr.; Marshall, Gerstein & Borun, 6300 Sears Tower, 233 South Wacker Drive, Chicago, IL 60606 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/31511 A2

(54) Title: MATERIALS AND METHODS TO MODULATE LIGAND BINDING/ENZYMATIC ACTIVITY OF α/β PROTEINS CONTAINING AN ALLOSTERIC REGULATORY SITE

(57) Abstract: Methods of modulating binding between an α/β protein and a binding partner are provided, along with methods of identifying modulators and their use.

MATERIALS AND METHODS TO MODULATE LIGAND BINDING/ENZYMATIC ACTIVITY OF α/β PROTEINS CONTAINING AN ALLOSTERIC REGULATORY SITE

CROSS REFERENCE TO RELATED APPLICATION

This application claims the benefit of U.S. Provisional Application Serial No. 60/239,750, filed October 12, 2000.

FIELD OF THE INVENTION

The present invention provides materials and methods to regulate binding activity of alpha/beta (α/β) molecules comprising an allosteric regulatory site.

BACKGROUND OF THE INVENTION

The alpha/beta (α/β) domain superfamily of proteins includes approximately ninety-seven families identified by specific fold structures. Proteins in the superfamily generally possess distinctive fold structures such as a TIM barrel, a horsehead fold or a beta-alpha-beta structure wherein a central beta sheet is surrounded by alpha helices, and is formed from multiple beta strand domains arranged in a parallel, anti-parallel or mixed orientation.

Many members of the superfamily, including proteins comprising an integrin I domain, von Willebrand factor comprising A domain structures, and various enzymes, have an open twisted beta sheet which gives rise to a fold in the protein's three dimensional structure. This fold is commonly referred to as a Rossmann fold, a Rossmann-like fold, or a dinucleotide binding fold. Many functionally diverse proteins contain Rossman folds, and these proteins can be identified using the SCOP, SMART, and CATH databases. A prototypic Rossmann fold is found at the site of NADP binding in glyceraldehyde-3-phosphate dehydrogenase.

Many Rossmann domains include a functional site on the "upper face" of the central beta sheet. This site in, for example, integrin I domains, Rho/Rac GTPases, and heterotrimeric GTPases, permits coordinated metal ion binding. In at least some integrin I domains, the bound metal ion forms a critical direct contact with a bound ligand and this site of metal ion binding has been designated the metal ion

dependent adhesion site (MIDAS). Metal ion binding sites in other proteins are also proximal to ligand binding, including, for example, GTP/GDP binding to GTPases, and cofactor (*i.e.*, NAD and FAD) binding to the bacterial protein ENR. Previous work has shown that for at least some proteins, including GTPases, LFA-1 [Huth, *et al.*, Proc. Natl. Acad. Sci. (USA) 97:5231-5236 (2000)], Mac-1 [Oxvig, *et al.*, Proc. Natl. Acad. Sci. (USA) 96:2215-20 (1999)] and Alpha2 [Emsley, *et al.*, Cell 101:47-56 (2000)], ligand binding in the MIDAS region requires a conformation change between the active and inactive state of the protein.

The integrin I domain structure has been characterized in detail. Among the integrins in which I domain structures have been identified, primary amino acid sequence comparison indicates that overall homology can vary widely among different integrin family members. Despite this divergence in homology, some residues are highly conserved in many integrins. Further, it has remained unclear whether the observed divergence in amino acid sequence homology gives rise to substantial differences in tertiary structure of the I domain within the individual subunits or the quaternary structure in the heterodimers.

The I domains for α_M [Lee *et al.*, Cell 80:631-638 (1995)], α_L [Qu *et al.*, Structure 4:931-942 (1996)], α_1 [Rich, J. Biol. Chem., 274:24906-24913 (1999)], and α_2 [Emsley *et al.*, J. Biol. Chem., 272:28512-28517 (1997)] have been crystallized, thereby permitting detailed analysis of previously speculated functional regions. The α_M crystalline structure clearly identified a Rossmann fold including a ligand-binding crevice formed along the top of the central, hydrophobic beta sheet, wherein the beta sheet is surrounded by multiple amphipathic α helices [Dickeson, *et al.*, Cell. Mol. Life. Sci. 54:556-566 (1998)]. Consistent with previous observations, crystalline I domains for both α_M and α_L have also been shown to include a MIDAS region.

General structural observations from the crystalline α_M I domain appear to correlate to the crystalline structure of α_L. These observations clearly indicate that α_L undergoes a conversion from an inactive to an active state before ligand binding can occur. This observation has been confirmed in NMR studies wherein ICAM-1 binding to the α_L I domain was shown to require positional perturbations of amino

acid residues in the α_L MIDAS region, as well as in a second region, still within the I domain but distal to the MIDAS region [Huth, *et al.*, Proc. Natl. Acad. Sci. (USA) 97:5231-5236 (2000)].

Site directed mutagenesis in this second region has indicated that residues therein are not part of the ICAM-1 binding site, *i.e.*, these residues do not interact directly with the ligand, but that these residues do, at least in part, play a role in regulating ICAM-1 binding. Amino acid residues that comprise this region have been designated the I domain allosteric site (IDAS) [Id.], and it is postulated that this region undergoes and/or induces a functionally relevant conformational shift that may be modulated by a small molecule. If the overall tertiary structure is conserved in the I or A domains of other proteins, such a site could provide an attractive target for modulating ligand binding for these proteins.

Furthermore, the crystal structure of the entire extracellular region of alphaVbeta3, an integrin, was recently reported [Cousin, Science, 293:1743-1746 (September 7, 2001)]. The crystal structure confirms predictions that the beta subunit of all integrins contains an I domain. Because this I domain has been implicated in regulating integrin function, it is an additional potential site for modulating ligand binding for these proteins. Identification of such regulatory regions provides means by which modulators, *i.e.*, agonists and antagonists, of ligand binding can be identified. Identification of such modulators provides candidate compounds that can provide protection against, and relief from, the myriad of pathological states associated with aberrant activity of α/β proteins.

Accordingly, there exists a need in the art to identify modes of modulating α/β proteins, which have a wide variety of functions and primary structures, in such a manner as to influence their biological activity.

SUMMARY OF THE INVENTION

In one aspect, the present invention provides methods of modulating binding interaction between a first molecule which is not LFA-1 or an I domain-containing fragment thereof, and a binding partner molecule, said first molecule comprising an α/β domain structure, said α/β structure comprising an allosteric

regulatory site, said method comprising the step of contacting said first molecule with an allosteric effector molecule that interacts with said allosteric regulatory site and promotes a conformation in a ligand binding domain of said α/β structure that modulates binding between said first molecule and said binding partner molecule. As used herein, the term "α/β structure" for a molecule refers to a general class of molecules that comprise a characteristic structure which is not necessarily indicative of, for example, molecules having multiple subunits which are designates as α and β subunits. This general class of molecules, however, can include molecules having multiple subunits which are designates as α and β subunits. The invention further provides methods of modulating binding interaction between a first molecule which is not LFA-1 or an I domain-containing fragment thereof, and a binding partner molecule, said first molecule comprising an α/β domain structure, said α/β structure comprising an allosteric regulatory site, said method comprising the step of contacting said first molecule with an allosteric effector molecule, said allosteric effector molecule comprising a diaryl compound, said diaryl compound interacting with said allosteric regulatory site and promoting a conformation in a ligand binding domain of said α/β structure that modulates binding between said first molecule and said binding partner molecule. In another aspect, the invention provides methods of modulating binding interaction between a first molecule which is not LFA-1 or an I domain-containing fragment thereof, and a binding partner molecule, said first molecule comprising an α/β domain structure, said α/β structure comprising an allosteric regulatory site, said method comprising the step of contacting said first molecule with an allosteric effector molecule, said allosteric effector molecule selected from the group consisting of diaryl sulfide compounds and diarylamide compounds, said allosteric effector molecule interacting with said allosteric regulatory site and promoting a conformation in a ligand binding domain of said α/β structure that modulates binding between said first molecule and said binding partner molecule.

In one embodiment, methods of the invention utilize a first molecule which comprises a Rossmann fold structure, said Rossmann fold structure comprising said allosteric regulatory site. As used herein, the term Rossmann fold structure encompasses Rossmann-like fold structures and dinucleotide fold structures, as is

known in the art. In the methods the Rossmann fold structure in the first molecule comprises a β sheet having β sheet strands positioned in a 321456 or 231456 orientation. Alternatively, the Rossmann fold structure in the first molecule comprises a β sheet having β sheet strands positioned in a 3214567 orientation. In another aspect, the Rossmann fold structure in said first molecule comprises a β sheet having β sheet strands positioned in a 32145 orientation. As used herein, the term orientation refers to the positioning of the individual strands of a β sheet in a parallel, antiparallel or mixed configuration. Preferably, methods employ a first molecule which comprises an I domain structure or an A domain structure.

The invention further provides methods of modulating binding interaction between a first molecule and a binding partner molecule, said first molecule having an amino acid sequence which exhibits less than about 90% identity to the LFA-1 I domain amino acid sequence set out in FIGURE 1, said first molecule comprising an α/β structure, said α/β domain structure comprising an allosteric regulatory site, said method comprising the step of contacting said first molecule with an allosteric effector molecule that interacts with said allosteric regulatory site and promotes a conformation in a ligand binding domain of said α/β structure that modulates binding between said first molecule and said binding partner molecule. The allosteric regulatory sites of the present invention include "I-like domains" or "IDAS-like domains," as well as IDAS domains. As used herein, the terms I-like domains and IDAS-like domains refer to regulatory sites discrete (i.e., distinguishable) from the MIDAS region (in MIDAS-containing molecules), and discrete (i.e., distinguishable) from ligand, substrate or co-factor binding sites, that do not necessarily include a complete I domain *per se*, but do undergo and/or induce a functionally relevant conformational shift that may be modulated by a small molecule to increase or decrease binding between a first molecule and a binding partner molecule. In another aspect, the invention provides methods of modulating binding interaction between a first molecule and a binding partner molecule, said first molecule having an amino acid sequence which exhibits less than about 90% identity to the LFA-1 I domain amino acid sequence set out in FIGURE 1, said first molecule comprising an α/β structure, said α/β domain structure comprising an allosteric

regulatory site, said method comprising the step of contacting said first molecule with an allosteric effector molecule, said allosteric effector molecule comprising a diaryl compound, said diaryl compound interacting with said allosteric regulatory site and promoting a conformation in a ligand binding domain of said α/β structure that modulates binding between said first molecule and said binding partner molecule. In still another aspect, the invention provides methods of modulating binding interaction between a first molecule and a binding partner molecule, said first molecule having an amino acid sequence which exhibits less than about 90% identity to the LFA-1 I domain amino acid sequence set out in FIGURE 1, said first molecule comprising an α/β domain structure, said α/β structure comprising an allosteric regulatory site, said method comprising the step of contacting said first molecule with an allosteric effector molecule, said allosteric effector molecule selected from the group consisting of diaryl sulfide compounds and diarylamide compounds, said allosteric effector molecule interacting with said allosteric regulatory site and promoting a conformation in a ligand binding domain of said α/β structure that modulates binding between said first molecule and said binding partner molecule. In a preferred embodiment, each of the methods the first molecule has an amino acid sequence that exhibits a percent identity with respect to the LFA-1 I domain amino acid sequence less than about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, or about 90%. In another aspect, the first molecule comprises a Rossmann fold structure, said Rossmann fold structure comprising an allosteric regulatory site and the first molecule has an amino acid sequence that exhibits a percent identity with respect to the LFA-1 I domain amino acid sequence less than about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, or about 90%. In another aspect, the methods of the invention utilize a first molecule wherein the Rossmann fold structure in said first molecule comprises a β sheet having β sheet strands positioned in a 321456 or 231456 orientation and the first molecule has an amino acid sequence that exhibits a percent identity with respect to the LFA-1 I domain amino acid sequence less than about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, or about 90%. In another aspect, the methods use a protein

wherein the Rossmann fold structure in said first molecule comprises a β sheet having β sheet strands positioned in a 3214567 orientation and the first molecule has an amino acid sequence that exhibits a percent identity with respect to the LFA-1 I domain amino acid sequence less than about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, or about 90%. In another aspect, the method utilize a first molecule with a Rossmann fold structure comprising a β sheet having β sheets strands positioned in a 32145 orientation, and the first molecule has an amino acid sequence that exhibits a percent identity with respect to the LFA-1 I domain amino acid sequence less than about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, 85%, or about 90%. Preferably, the first molecule comprises an I domain structure and the first molecule has an amino acid sequence that exhibits a percent identity with respect to the LFA-1 I domain amino acid sequence less than about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%,or about 90%. In another preferred embodiment, the first molecule comprises an A domain structure and the first molecule has an amino acid sequence that exhibits a percent identity with respect to the LFA-1 I domain amino acid sequence less than about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, or about 90%.

In methods of the invention, the modulator promotes a conformation in the ligand binding domain of said first molecule that increases binding between said first molecule and said binding partner molecule, and in one aspect, the increase in binding between the first molecule and the second molecule results in increased enzymatic activity of the first molecule. In another embodiment, the modulator promotes a conformation in the ligand binding domain of said first molecule that decreases binding between said first molecule and said binding partner molecule and the decrease in binding between the first molecule and the second molecule results in decreased enzymatic activity of the first molecule.

Methods include use of a first molecule selected from the group consisting of the proteins set forth in Table 1 as well as other proteins which comprise I or A domains, G proteins, heterotrimeric G proteins, and tubulin GTPase.

Preferably, methods of the invention utilize a first molecule selected from the group consisting of the proteins set forth in Table 1. In one aspect, the first molecules is a eukaryotic molecule. Preferably, the first molecule is a human molecule. In another aspect, the first molecule is a prokaryotic molecule. In one embodiment, the first molecule is a bacterial molecule.

More preferably, the first molecule is selected from the group consisting of $\alpha_M\beta_2$, complement protein C2, complement protein Factor B, $\alpha_E\beta_7$, $\alpha_4\beta_7$, $\alpha_V\beta_3$, $\alpha_4\beta_1$, $\alpha_d\beta_2$,von Willebrand factor, Rac-1, HPPK, ftsZ, and ENR. In methods wherein the first molecule is $\alpha_M\beta_2$ and the binding partner protein is fibrinogen; the first molecule is $\alpha_M\beta_2$ and the binding partner protein is iC3b; the first molecule is $\alpha_E\beta_7$ and the binding partner protein is E-cadherin; the first molecule is $\alpha_4\beta_7$ and the binding partner protein is MadCAM-1; the first molecule is $\alpha_V\beta_3$ and the binding partner protein is vitronectin; the first molecule is $\alpha_4\beta_1$ and the binding partner protein is VCAM; the first molecule is $\alpha_d\beta_2$ and the binding partner protein is VCAM; the first molecule is von Willebrand factor and the binding partner protein is gpIb; the first molecule is complement protein C2 and the binding partner protein is complement protein C4b; the first molecule is complement protein Factor B and the binding partner protein is complement protein C3b; the first molecule is Rac-1 and the binding partner is GTP; the first molecule is HPPK and the binding partner is ATP or HMDP; the first molecule is ftsZ and the binding partner is GTP; and the first molecule is ENR and the binding partner is NADH.

DETAILED DESCRIPTION OF THE INVENTION

In one aspect, the present invention provides methods of modulating binding interaction between a first molecule which is not LFA-1 or an I domain-containing fragment or mimetics thereof, and a binding partner molecule, said first molecule comprising an α/β structure, said α/β structure comprising an allosteric regulatory site, said method comprising the step of contacting said first molecule with an allosteric effector molecule that interacts with said allosteric regulatory site and promotes a conformation in a ligand binding domain of said α/β structure that modulates binding between said first molecule and said binding partner molecule. As

used herein, "binding partner molecules" includes ligands, substrates and cofactor, the binding of which is required to effect one or more biological activity of the first molecule. An I domain fragment of LFA-1 is a polypeptide portion or fragment (*i.e.*, a polypeptide that is less than full length LFA-1 as set out in FIGURE 2) of LFA-1 that comprises (i) the I domain of LFA-1, or (ii) a portion of the LFA-1 I domain that maintains biologically active features of the LFA-1 I domain. Synthetic mimetics of the LFA-1 I domain, including peptidomimetics which replicate or affect one or more biological activities of the LFA-1 I domain, are also included in this definition. The α/β superfamily of proteins includes those proteins having an beta-alpha-beta structure wherein a central beta sheet domain is flanked on both sides of the sheet by one or more alpha helix domains.

In another aspect, the present invention provides methods of modulating binding interaction between a first molecule which is not LFA-1 or an I domain-containing fragment or mimetics thereof, and a binding partner molecule, said first molecule comprising a Rossmann fold structure, said Rossmann fold structure comprising an allosteric regulatory site, said method comprising the step of contacting said first molecule with an allosteric effector molecule that interacts with said allosteric regulatory site and promotes a conformation in a ligand binding domain of said Rossmann fold structure that modulates binding between said first molecule and said binding partner molecule. A Rossmann fold structure in a protein comprises a beta sheet structure wherein individual beta sheet domains of the protein are positioned in either parallel, antiparallel, or mixed orientations. In preferred aspects of the present invention, the beta sheet of the first molecule is comprised of individual beta sheet strands. Numerical designations for the individual beta sheet strands are assigned according to their position in the primary amino acid sequence of the first protein, with the first beta sheet strand being that one closest to the amino terminus of the protein sequence. Rossmann fold structures are further characterized by the presence of a ligand binding fold, pocket, or site in the three dimensional structure of the beta sheet that is generally positioned at the "top" of the beta sheet structure.

In another aspect, the present invention provides methods of modulating binding interaction between a first molecule which is not LFA-1 or an I

domain-containing fragment or mimetic thereof, and a binding partner molecule, said first molecule comprising a Rossmann fold structure, said Rossmann fold structure comprising a β sheet having β strands positioned in a 321456 or 231456 orientation and an allosteric regulatory site, said method comprising the step of contacting said first molecule with an allosteric effector molecule that interacts with said allosteric regulatory site and promotes a conformation in a ligand binding domain of said Rossmann fold structure that modulates binding between said first molecule and said binding partner molecule. In another aspect, the present invention provides methods of modulating binding interaction between a first molecule which is not LFA-1 or an I domain-containing fragment or mimetic thereof, and a binding partner molecule, said first molecule comprising a Rossmann fold structure, said Rossmann fold structure comprising a β sheet having β strands positioned in a 3214567 orientation and an allosteric regulatory site, said method comprising the step of contacting said first molecule with an allosteric effector molecule that interacts with said allosteric regulatory site and promotes a conformation in a ligand binding domain of said Rossmann fold structure that modulates binding between said first molecule and said binding partner molecule. The present invention also provides methods of modulating binding interaction between a first molecule which is not LFA-1 or an I domain-containing fragment or mimetic thereof, and a binding partner molecule, said first molecule comprising a Rossmann fold structure, said Rossmann fold structure comprising a β sheet having β strands positioned in a 32145 orientation and an allosteric regulatory site, said method comprising the step of contacting said first molecule with an allosteric effector molecule that interacts with said allosteric regulatory site and promotes a conformation in a ligand binding domain of said Rossmann fold structure that modulates binding between said first molecule and said binding partner molecule. Numerical designations for individual beta sheets in the first molecule are as described above.

In another aspect, the present invention provides methods of modulating binding interaction between a first molecule which is not LFA-1 or an I domain-containing fragment or mimetic thereof, and a binding partner molecule, said first molecule comprising an I domain structure, said I domain structure comprising

an allosteric regulatory site, said method comprising the step of contacting said first molecule with an allosteric effector molecule that interacts with said allosteric regulatory site and promotes a conformation in a ligand binding domain of said I domain structure that modulates binding between said first molecule and said binding partner molecule. I domain structures are known in the art to comprise approximately 200 amino acids as exemplified by the domains identified in a number of integrins [See Dickeson, *et al.*, Cell. Mol. Life Sci. 54:556-566 (1998)].

The present invention also provides methods of modulating binding interaction between a first molecule which is not LFA-1 or an I domain-containing fragment thereof, and a binding partner molecule, said first molecule comprising an A domain structure, said A domain structure comprising an allosteric regulatory site, said method comprising the step of contacting said first molecule with an allosteric effector molecule that interacts with said allosteric regulatory site and promotes a conformation in a ligand binding domain of said A domain structure that modulates binding between said first molecule and said binding partner molecule. A domain motifs are known in the art to share homology with I domains and are exemplified by the domains found in von Willebrand factor.

The present invention also provides methods of modulating binding interaction between a first molecule and a binding partner molecule, said first molecule having an amino acid sequence which exhibits less than about 90% identity to the LFA-1 I domain amino acid sequence [set out in FIGURE 1], said first molecule comprising an α/β structure, said α/β structure comprising an allosteric regulatory site, said method comprising the step of contacting said first molecule with an allosteric effector molecule that interacts with said allosteric regulatory site and promotes a conformation in a ligand binding domain of said α/β structure that modulates binding between said first molecule and said binding partner molecule. Identity as used herein can be calculated using basic BLAST analysis using default parameters. Values for percent identity reflect one-to-one correspondence between amino acid residues across the entire LFA-1 sequence I domain as set out in FIGURE 1 and a region of amino acid residues of the same or similar length in the first molecule. In another embodiment of the method, the first molecule has an amino acid

sequence that exhibits a percent identity with respect to the LFA-1 I domain amino acid sequence of less than about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, or about 90%.

In still another aspect, the present invention provides methods of modulating binding interaction between a first molecule and a binding partner molecule, said first molecule having an amino acid sequence which exhibits less than about 90% identity to the LFA-1 I domain amino acid sequence [set out in FIGURE 1], said first molecule comprising a Rossmann fold structure, said Rossmann fold structure comprising an allosteric regulatory site, said method comprising the step of contacting said first molecule with an allosteric effector molecule that interacts with said allosteric regulatory site and promotes a conformation in a ligand binding domain of said Rossmann fold structure that modulates binding between said first molecule and said binding partner molecule. In alternative embodiments of the method, the first molecule has an amino acid sequence that exhibits a percent identity with respect to the LFA-1 I domain amino acid sequence of less than about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, or about 90%.

In another aspect, the present invention provides methods of modulating binding interaction between a first molecule and a binding partner molecule, said first molecule having an amino acid sequence which exhibits less than about 90% identity to the LFA-1 I domain amino acid sequence [set out in FIGURE 1], said first molecule comprising a Rossmann fold structure with β sheets strands positioned in a 321456 or 231456 orientation and an allosteric regulatory site, said method comprising the step of contacting said first molecule with an allosteric effector molecule that interacts with said allosteric regulatory site and promotes a conformation in a ligand binding domain of said Rossmann fold structure that modulates binding between said first molecule and said binding partner molecule. In alternative embodiments of the method, the first molecule has an amino acid sequence that exhibits a percent identity with respect to the LFA-1 I domain amino acid sequence of less than about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, or about 90%.

In still another aspect, the present invention provides methods of modulating binding interaction between a first molecule and a binding partner molecule, said first molecule having an amino acid sequence which exhibits less than about 90% identity to the LFA-1 I domain amino acid sequence [set out in FIGURE 1], said first molecule comprising a Rossmann fold structure, said Rossmann fold structure with β sheet strands positioned in a 3214567 orientation and an allosteric regulatory site, said method comprising the step of contacting said first molecule with an allosteric effector molecule that interacts with said allosteric regulatory site and promotes a conformation in a ligand binding domain of said Rossmann fold structure that modulates binding between said first molecule and said binding partner molecule. In alternative embodiments of the method, the first molecule has an amino acid sequence that exhibits a percent identity with respect to the LFA-1 I domain amino acid sequence of less than about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, or about 90%.

The present invention also provides methods of modulating binding interaction between a first molecule and a binding partner molecule, said first molecule having an amino acid sequence which exhibits less than about 90% identity to the LFA-1 I domain amino acid sequence [set out in FIGURE 1], said first molecule comprising a Rossmann fold structure β sheet strands positioned in a 32145 orientation and an allosteric regulatory site, said method comprising the step of contacting said first molecule with an allosteric effector molecule that interacts with said allosteric regulatory site and promotes a conformation in a ligand binding domain of said Rossmann fold structure that modulates binding between said first molecule and said binding partner molecule. In alternative embodiments of the method, the first molecule has an amino acid sequence that exhibits a percent identity with respect to the LFA-1 I domain amino acid sequence of less than about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, or about 90%.

The present invention further provides methods of modulating binding interaction between a first molecule and a binding partner molecule, said first molecule having an amino acid sequence which exhibits less than about 90% identity

to the LFA-1 I domain amino acid sequence [set out in FIGURE 1], said first molecule comprising an I domain structure, said I domain structure comprising an allosteric regulatory site, said method comprising the step of contacting said first molecule with an allosteric effector molecule that interacts with said allosteric regulatory site and promotes a conformation in a ligand binding domain of said I domain structure that modulates binding between said first molecule and said binding partner molecule. In alternative embodiments of the method, the first molecule has an amino acid sequence that exhibits a percent identity with respect to the LFA-1 I domain amino acid sequence of less than about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, or about 90%.

In another aspect, the present invention provides methods of modulating binding interaction between a first molecule and a binding partner molecule, said first molecule having an amino acid sequence which exhibits less than about 90% identity to the LFA-1 I domain amino acid sequence [set out in FIGURE 1], said first molecule comprising an A domain structure, said A domain structure comprising an allosteric regulatory site, said method comprising the step of contacting said first molecule with an allosteric effector molecule that interacts with said allosteric regulatory site and promotes a conformation in a ligand binding domain of said A domain structure that modulates binding between said first molecule and said binding partner molecule. In alternative embodiments of the method, the first molecule has an amino acid sequence that exhibits a percent identity with respect to the LFA-1 I domain amino acid sequence of less than about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, or about 90%.

In each of the methods of the present invention, the modulator promotes a conformation in the ligand binding domain of said first molecule that increases binding between said first molecule and said binding partner molecule. Alternatively, the modulator promotes a conformation in the ligand binding domain of said first molecule that decreases binding between said first molecule and said binding partner molecule. Preferably, the methods include a first molecule selected from the

group consisting of the molecules set out in Table 1 or otherwise described herein. Most preferably, methods utilize a first molecule selected from the group consisting of $\alpha_M\beta_2$, complement protein C2, complement protein Factor B, $\alpha_E\beta_7$, $\alpha_4\beta_7$, $\alpha_V\beta_3$, $\alpha_4\beta_1$,$\alpha_d\beta_2$ von Willebrand factor, Rac-1, HPPK, ftsZ, and ENR. Furthermore, preferably, the methods and compositions of the present invention use a modulator that is a diaryl compound. More preferably, the methods and compositions of the present invention use a modulator that is selected from diaryl sulfide compounds and diarylamide compounds. Most preferably, the methods and compositions of the present invention use a modulator that is a diaryl sulfide compound.

In methods wherein the first molecule is $\alpha_M\beta_2$, the preferred binding partner protein is fibrinogen, and a preferred modulator is selected from the group consisting of Cmpd S, Cmpd R, Cmpd N, Cmpd O, Cmpd P, Cmpd Q, Cmpd L, Cmpd V, Cmpd F, Cmpd AA, and Cmpd AC as set out in Table 2. In methods wherein the first molecule is $\alpha_M\beta_2$, an alternative preferred binding partner protein is iC3b and a preferred modulator is selected from the group consisting of Cmpd H, Cmpd I and Cmpd C. In methods wherein the first molecule is $\alpha_E\beta_7$, the preferred binding partner protein is E-cadherin and a preferred modulator is selected from the compounds set out in Table 2 herein. In methods wherein the first molecule is $\alpha_4\beta_7$, the preferred binding partner protein is MAdCAM-1. In methods wherein the first molecule is $\alpha_V\beta_3$, the preferred binding partner protein is vitronectin. In methods wherein the first molecule is $\alpha_4\beta_1$, the preferred binding partner protein is VCAM. In methods wherein the first molecule is $\alpha_d\beta_2$, the preferred binding partner protein is VCAM. In methods wherein the first molecule is von Willebrand factor, the preferred binding partner protein is gpIb. In methods wherein the first molecule is complement protein C2 , the preferred binding partner protein is complement protein C4b. In methods wherein the first molecule is complement protein Factor B, the preferred binding partner protein is complement protein C3b. In methods wherein the first molecule is either $\alpha_1\beta_1$, $\alpha_2\beta_1$, $\alpha_{11}\beta_1$, the preferred binding partner is collagen. In methods wherein the first molecule is $\alpha_2\beta_1$, the preferred binding partner is collagen and a preferred modulator is selected from the group of compounds set out in Table 2 herein. In methods wherein the first molecule is Rac-1, the preferred binding partner

is GDP/GTP and a preferred modulator GTP. In methods wherein the first molecule is HPPK, the preferred binding partner is ATP or HMDP. In methods wherein the first molecule is ftsZ, the preferred binding partner is GTP. In methods wherein the first molecule is ENR, the preferred binding partner is NADH.

Methods of the present invention include those wherein the first molecule, the binding partner molecule or both are isolated proteins, or binding fragments thereof, obtained from natural sources or from cells modified to express the molecules as heterologous proteins. The methods also embrace use of the first molecule, or a binding fragment thereof, the binding partner molecule, or a binding fragment thereof, both which are expressed on the surface of cells which express the molecules as homologous proteins or on the surface of cells which have been modified to express heterologous proteins. *In vivo* and *in vitro* methods are contemplated.

In vivo methods are expected to alleviate and/or prevent pathological states which arise from aberrant binding activity between the first molecule and the binding partner molecule. For example, indications associated with inappropriate complement activation for which methods of the present invention are expected to alleviate or prevent include: (i) diseases involving antibody/complement deposition which includes systemic lupus erythematosus (SLE), Goodpasture's disease, rheumatoid arthritis, myasthenia gravis, autoimmune hemolytic anemia, autoimmune thrombocytopenic purpura, and Rasmussen's encephalitis; (ii) diseases involving ischemia-reperfusion injury, including stroke, myocardial infarction, cardiac pulmonary bypass, acute hypovolemic disease, renal failure, and allotransplantation; (iii) central nervous system pathologies such as Alzheimer's disease and multiple sclerosis; and (iv) miscellaneous indications such as trauma, chemical or thermal injury, and xenotransplantation.

Likewise, inhibitors of alpha 1, alpha 2, and alpha 11 are also expected to be useful for treating cancer. During metastasis, tumor cells must pass through the extracellular matrix prior to intravasation and following extravasation. Migration through these regions is dependent on integrin activity. In addition, it has been shown that blocking of α_1 or α_2 activity with monoclonal antibodies [Locher *et al.*, Mol. Biol.

Cell. 10:271-282 (1999)] or removal of α_1 activity in a knockout mouse [Pozzi, *et al.*, Proc. Natl. Acad. Sci. (USA) 97:2202-2207 (2000)] results in changes in matrix metalloproteinase (MMP) levels. MMPs are extracellular matrix-degrading enzymes which have been proposed to play a role in a variety of types of cancer. [For a review, see Nelson, et. al., J. Clin. Oncol. 18:1135-1149 (2000)]. Inhibitors of MMPs are currently being tested for clinical utility in treating many types of cancer. To date, MMP inhibitors have not been as effective in human trials as in animal models. Modulating MMP expression by inhibiting integrin activity can prove to be more effective by differentially modulating different MMP levels and by specifically targeting this MMP modulation to α_1, α_2, or α_{11} expressing cells.

More particularly, it has been demonstrated that alpha 11 is expressed on foamy macrophages in atherosclerotic plaques as well as in a subset of macrophages in synovium from a patient with rheumatoid arthritis. No expression has been seen in non-activated monocyte derived macrophages. Inhibitors of alpha 11/ligand binding interactions could therefore be useful for reducing migration and/or signaling events of macrophages that are associated with different inflammatory processes. Accordingly, alpha 11 inhibitors could represent useful therapeutics for treating inflammatory diseases, including atherosclerosis and rheumatoid arthritis.

Similarly, alpha 1 and alpha 2 integrins have been shown to be upregulated on certain cells (including T cells and monocytes) following stimulation. It has also been demonstrated that blocking interactions between alpha 1 or alpha 2 and their ligands using monoclonal antibodies inhibited inflammatory responses in mouse models of delayed-type hypersensitivity, contact hypersensitivity and arthritis [deFougerolles *et. al. J. Clin. Invest. 105*:721-729(2000)]. Antagonists of alpha1 and alpha 2 may inhibit inflammation through a variety of mechanisms including inhibiting cell migration, cell proliferation and the production of inflammatory mediators such as matrix metalloproteinase 3, tumor necrosis factor alpha and interleukin-1. Accordingly, small molecule inhibitors or antagonists of alpha1 and alpha2 associations (ligand binding), *i.e.*, allosteric effector molecules, could be useful for the treatment of inflammatory diseases such as arthritis, fibrotic diseases and cancer.

Fibrotic disease states are characterized by the excessive production of fibrous extracellular matrix by certain cell types that are inappropriately activated. It is believed that the mechanism of fibrous extracellular matrix formation involves, at least in part, α/β protein activity. Accordingly, by inhibiting α/β proteins, the present invention provides methods and compositions for the treatment and prevention of various fibrotic disease states, including scleroderma (morphea, generalized morphea, linear scleroderma), keloids, hypertrophic scar, nodular fasciitis, eosinophilic fasciitis, Dupuytren's contracture, kidney fibrosis, pulmonary fibrosis, chemotherapy / radiation induced lung fibrosis, atherosclerotic plaques, inflammatory bowel disease, Crohn's disease, arthritic joints, invasive breast carcinoma desmosplasis, dermatofibromas, endothelial cell expression, angiolipoma, angioleiomyoma, sarcoidosis, cirrhosis, idiopathic interstitial lung disease, idiopathic pulmonary fibrosis (4 pathologic types), collagen vascular disease associated lung syndromes, cryptogenic organizing pneumonia, Goodpasture's syndrome, Wegener's granulomatosis, eosinophilic granuloma, iatrogenic lung disease, pneumoconioses (asbestosis, silicosis), hypersensitivity pneumonitides (farmer's lung, bird fancier's lung, etc.), interstitial pulmonary fibrosis, chemical pneumonitis, hypersensitivity pneumonitis and the like.

With respect to bacterial proteins, ENR is already a target for anti-tuberculosis drugs and a target of the broad spectrum biocide triclosan. Small molecules would therefore be useful in drug resistant tuberculosis. Moreover, the activity spectrum of ENR and DapB inhibitors would be useful as Gram negative inhibitors. Furthermore, because ERA-GTPase is highly conserved among bacteria, inhibitors would be useful against a broad spectrum of bacteria, depending on permeability. In addition, inhibitors of the various bacterial proteins would be useful for treating bacterial diseases involving Gram negative bacteria and infections with undefined bacterial pathogens.

Other chemotherapeutics, such as sulfonamides, inhibit bacterial growth by antagonizing the *de novo* folate biosynthetic pathway [Mandell and Petri, *Sulfonamides, Trimethoprim-sulfamethoxazole, Quinolones, and Agents for Urinary Tract Infections, in* The Pharmacological Basis of Therapeutics (Goodman and Gilman eds., 1996)]. The primary goal of anti-folate therapy is to deplete the

intracellular pools of reduced folate, resulting in the inhibition of DNA replication due to insufficient levels of thymidine [Hitchings and Baccanari, *Design and Synthesis of Folate Antagonists as Antimicrobial Agents, in* Folate Antagonists as Therapeutic Agents (1984)].

The enzyme 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase (HPPK) catalyzes the transfer of pyrophosphate from ATP to 6-hydroxy-7,8-dihydropterin (HMDP) in the *de novo* folate biosynthetic pathway [Richey and Brown, *J. Biol. Chem.*, 244:1582-1592 (1969)]. HPPK is expressed in both Gram positive and Gram negative bacteria, fungi, and protozoa, but not in higher eukaryotes, and represents an important target for the development of antibiotics with anti-folate activity. By inhibiting HPPK, the present invention can provide methods and compositions for the treatment and prevention of various bacterial and fungal infections.

FtsZ is the product of an essential bacterial gene that is involved in cell division. FtsZ binds and hydrolyzes GTP, and when bound to GTP it forms long, linear polymers. The GTP-dependent polymerization of ftsZ is related to its function in bacterial cell division. During septation, ftsZ forms a ring to define the plane of cell division. Cells lacking ftsZ can not undergo septation, do not divide and die. FtsZ is highly conserved (approximately 60%) throughout the bacterial kingdom. Accordingly, by inhibiting ftsZ, the compositions and methods of the present invention provide broad-spectrum antibiotics. The atomic structure of ftsZ shows that it is an alpha/beta protein [Nogales *et al.*, (1998) Nature Structural Biology 5:451-458].

Modulators of vWF binding are useful in treatment of thrombotic vascular diseases, such as myocardial infarction (MI) and thrombotic stroke. Acute administration of a vWF A1-domain binding antagonist can reduce the risk of coronary vascular occlusion in high risk patients such as those with unstable angina, or following PTCA or stent placement. Several gpIIb/IIIa antagonists have recently been approved for clinical use in these settings (ReoPro®, Itrafiban, sibrafiban). While these agents are effective, their use is accompanied by bleeding, thus limiting their effective dose. If the bleeding side effects of an A1-domain inhibitor are limited,

it can be used chronically in individuals at risk for vascular occlusion. These individuals include patients with angina, claudication, and those with a history of MI or stroke. Abnormalities of vWF metabolism are the cause of the occlusive thrombus in thrombotic thrombocytopenic purpura, suggesting A1 domain inhibitors may also be useful in this setting.

Rac1, Rac2 and Rac3 are members of the Ras superfamily of small molecular weight (approximately 22-25kDa) GTPases, many of which are α/β proteins [Edwards and Perkins, FEBS Lett 358:283 (1995); De Vos *et al.*, Science 239:888 (1988); Worthylake *et al.*, Nature 408:682 (2000)]. Primary amino acid sequence comparison indicates that the overall homology of the Rac proteins is about 88 to about 92 percent identical. It is known that Rac1 and Rac2 proteins play a crucial role in cell survival, proliferation, metastasis and reactive oxygen species (ROS) production [Symons, *Curr. Opin. in Biotech.*, 6:668 (1995); and, Scita, *EMBO J.*, 19(11):2393 (2000)]. Due to the importance of Rac proteins in the control of cell proliferation, antagonists of the Rac guanine nucleotide exchange reaction and, in particular, small molecules that interfere with the exchange of GDP for GTP of Rac1 in the presence of Tiam1, are of considerable interest for the methods and compositions of the present invention.

In view of the indications described above, the present invention further provides methods for alleviating or preventing a condition arising from aberrant binding between a first molecule that is not LFA-1 or an I domain fragment thereof and a binding partner molecule, wherein said first molecule is an α/β protein selected from the group of proteins set forth in Table 1 , said method comprising the steps of administering to an individual in need thereof an effective amount of a modulator of binding between said first molecule and said binding partner molecule. As used herein, the term effective amount refers to the administration of an amount of a modulator sufficient to achieve its intended purpose. More specifically, a "therapeutically effective amount" refers to an amount effective to treat or to prevent development of, or to alleviate the existing symptoms of, the subject being treated. Determination of the effective amounts is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein.

In one aspect, the present invention provides methods of treatment wherein the α/β protein comprises a Rossmann fold. In another aspect, methods of treatment are provided wherein the Rossmann fold in the targeted protein includes five, six or seven β strands which makeup the central β sheet structure. When the Rossmann fold comprises five β strands, it is preferred that the positioning of the individual strands is 32145 as defined above. When the Rossmann fold comprises six β strands, it is preferred that the positioning of the individual strands is 321456 or 231456 as defined above. When the Rossmann fold comprises seven β strands, it is preferred that the positioning of the individual strands is 3214567 as defined above. Methods of treatment the present invention include those wherein the first molecule exhibits less than about 90% amino acid sequence identity with the I domain amino acid sequence of LFA-1 as set out in FIGURE 1. Preferably, the first molecule will have a percent amino acid sequence identity with the I domain of LFA-1 less than about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, or about 90%. Sequence identity for purposes of this aspect of the present invention is calculated using, for example, basic BLAST search analysis with default parameters.

The present invention also provides methods for identifying a modulator of binding between a first molecule that is not LFA-1 or an I domain fragment thereof and a binding partner molecule, wherein said first molecule is an α/β protein selected from the group of proteins set forth in Table 1, said method comprising the steps of measuring binding between the first molecule and the binding partner molecule in the presence and absence of a test compound, and identifying the test compound as a modulator of binding when a change in binding between the first molecule and the binding partner molecule is detected in the presence of the test compound as compared to binding in the absence of the test compound. In one aspect, the present invention provides methods wherein the α/β protein comprises a Rossmann fold. In another aspect, methods are provided wherein the Rossmann fold in the targeted protein includes five, six or seven β strands which makeup the central β sheet structure. When the Rossmann fold comprises five β strands, it is preferred that the positioning of the individual strands is 32145 as defined above. When the

Rossmann fold comprises six β strands, it is preferred that the positioning of the individual strands is 321456 231456 as defined above. When the Rossmann fold comprises seven β strands, it is preferred that the positioning of the individual strands is 3214567 as defined above. Methods of the present invention include those wherein the first molecule exhibits less than about 90% amino acid sequence identity with the I domain amino acid sequence of LFA-1 as set out in FIGURE 1. Preferably, the first molecule will have a percent amino acid sequence identity with the I domain of LFA-1 less than about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, or about 90%. Sequence identity for purposes of this aspect of the present invention is calculated using, for example, basic BLAST search analysis with default parameters.

The present invention also provides modulators of binding between a first molecule that is not LFA-1 or an I domain fragment thereof and a binding partner molecule, wherein said first molecule is an α/β protein selected from the group of proteins set forth in Table 1. In one aspect, the modulators are those that affect binding of an α/β protein which comprises a Rossmann fold. In another aspect, modulators are provided which affect binding when the Rossmann fold in the targeted protein includes five, six or seven β strands which makeup the central β sheet structure. When the Rossmann fold comprises five β strands, it is preferred that the positioning of the individual strands is 32145 as defined above. When the Rossmann fold comprises six β strands, it is preferred that the positioning of the individual strands is 321456 or 231456 as defined above. When the Rossmann fold comprises seven β strands, it is preferred that the positioning of the individual strands is 3214567 as defined above. Modulators are also provided for a first molecule which exhibits less than about 90% amino acid sequence identity with the I domain amino acid sequence of LFA-1 as set out in FIGURE 1. Preferably, the first molecule will have a percent amino acid sequence identity with the I domain of LFA-1 less than about 40%, about 45%, about 50%, about 55%, about 60%, about 65%, about 70%, about 75%, about 80%, about 85%, or about 90%. Sequence identity for purposes of this aspect of the present invention is calculated using, for example, basic BLAST search analysis with default parameters.

The present invention also provides compositions comprising a modulator. Preferred compositions are pharmaceutical compositions. The pharmaceutical compositions of the present invention comprise one or more modulators of the present invention, preferably further comprising a pharmaceutically acceptable carrier or diluent. The term "pharmaceutically acceptable carrier" as used herein refers to compounds suitable for use in contact with recipient animals, preferably mammals, and more preferably humans, and having a toxicity, irritation, or allergic response commensurate with a reasonable benefit/risk ratio, and effective for their intended use.

The present invention also provides modulators which exist in a prodrug form. The term "prodrug" as used herein refers to compounds which are rapidly transformed *in vivo* to the parent, or active modulator, compound , for example, by hydrolysis. A thorough discussion is provided in Higuchi, *et al.*, Prodrugs as Novel Delivery Systems, vol. 14 of the A.C.S.D. Symposium Series, and in Roche (ed), Bioreversible Carriers in Drug Design, American Pharmaceutical Association and Pergamon Press, 1987, both of which are incorporated herein by reference. Prodrug design is discussed generally in Hardma, *et al.*, (Eds), Goodman & Gilman's The Pharmacological Basis of Therapeutics, Ninth Edition, New York, New York (1996), pp. 11-16. Briefly, administration of a drug is followed by elimination from the body or some biotransformation whereby biological activity of the drug is reduced or eliminated. Alternatively, a biotransformation process may lead to a metabolic by-product which is itself more active or equally active as compared to the drug initially administered. Increased understanding of these biotransformation processes permits the design of so-called "prodrugs" which, following a biotransformation, become more physiologically active in an altered state. Prodrugs are therefore pharmacologically inactive compounds which are converted to biologically active metabolites. In some forms, prodrugs are rendered pharmacologically active through hydrolysis of, for example, an ester or amide linkage, often times introducing or exposing a functional group on the prodrug. The thus modified drug may also react with an endogenous compound to form a water

soluble conjugate which further increases pharmacological properties of the compound, for example, as a result of increased circulatory half-life.

As another alternative, prodrugs can be designed to undergo covalent modification on a functional group with, for example, glucuronic acid sulfate, glutathione, amino acids, or acetate. The resulting conjugate may be inactivated and excreted in the urine, or rendered more potent than the parent compound. High molecular weight conjugates may also be excreted into the bile, subjected to enzymatic cleavage, and released back into circulation, thereby effectively increasing the biological half-life of the originally administered compound.

Compounds of the present invention may exist as stereoisomers where asymmetric or chiral centers are present. Stereoisomers are designated by either "S" or "R" depending on the arrangement of substituents around a chiral carbon atom. Mixtures of stereoisomers are contemplated by the present invention. Stereoisomers include enantiomers, diastereomers, and mixtures thereof. Individual stereoisomers of compounds of the present invention can be prepared synthetically from commercially available starting materials which contain asymmetric or chiral centers or by preparation of racemic mixtures followed by separation or resolution techniques well known in the art. Methods of resolution include (1) attachment of a mixture of enantiomers to a chiral auxiliary, separation of the resulting mixture by recrystallization or chromatography, and liberation of the optically pure product from the auxiliary; (2) salt formation employing an optically active resolving agent, and (3) direct separation of the mixture of optical enantiomers on chiral chromatographic columns.

The pharmaceutical compositions of the present invention can be administered to humans and other animals by any suitable route. For example, the compositions can be administered orally, rectally, parenterally, intracisternally, intravaginally, intraperitoneally, topically (as by powders, ointments, or drops), bucally, or nasally. The term "parenteral" administration as used herein refers to modes of administration which include intravenous, intraarterial, intramuscular, intraperitoneal, intrasternal, intrathecal, subcutaneous and intraarticular injection and infusion.

Pharmaceutical compositions of this present invention for parenteral injection comprise pharmaceutically-acceptable sterile aqueous or nonaqueous solutions, dispersions, suspensions or emulsions as well as sterile powders for reconstitution into sterile injectable solutions or dispersions just prior to use. Examples of suitable aqueous and nonaqueous carriers, diluents, solvents or vehicles include water, ethanol, polyols (such as glycerol, propylene glycol, polyethylene glycol, and the like), and suitable mixtures thereof, vegetable oils (such as olive oils), and injectable organic esters such as ethyl oleate. Proper fluidity can be maintained, for example, by the use of coating materials such as lecithin, by the maintenance of the required particle size, in the case of dispersions, and by the use of surfactants.

These compositions may also contain adjuvants such as preservatives, wetting agents, emulsifying agents, and dispersing agents. Prevention of the action of microorganisms may be ensured by the inclusion of various antibacterial and antifungal agents, for example, paraben, chlorobutanol, phenol sorbic acid, and the like. It may also be desirable to include isotonic agents such as sugars, sodium chloride, and the like. Prolonged absorption of the injectable pharmaceutical form may be brought about by the inclusion of agents which delay absorption such as aluminum monostearate and gelatin.

In some cases, in order to prolong the effect of the drug, it is desirable to slow the absorption of the drug from subcutaneous or intramuscular injection. This result may be accomplished by the use of a liquid suspension of crystalline or amorphous materials with poor water solubility. The rate of absorption of the drug then depends upon its rate of dissolution, which in turn may depend upon crystal size and crystalline form. Alternatively, delayed absorption of a parenterally administered drug from is accomplished by dissolving or suspending the drug in an oil vehicle.

Injectable depot forms are made by forming microencapsule matrices of the drug in biodegradable polymers such a polylactide-polyglycolide. Depending upon the ratio of drug to polymer and the nature of the particular polymer employed, the rate of drug release can be controlled. Examples of other biodegradable polymers include poly(orthoesters) and poly(anhydrides). Depot injectable formulations are

also prepared by entrapping the drug in liposomes or microemulsions which are compatible with body tissue.

The injectable formulations can be sterilized, for example, by filtration through a bacterial- or viral-retaining filter, or by incorporating sterilizing agents in the form of sterile solid compositions which can be dissolved or dispersed in sterile water or other sterile injectable medium just prior to use.

Solid dosage forms for oral administration include capsules, tablets, pills, powders, and granules. In such solid dosage forms, the active compound is mixed with a least one inert, pharmaceutically-acceptable excipient or carrier such as sodium citrate or dicalcium phosphate and/or (a) fillers or extenders such as starches, lactose, sucrose, glucose, mannitol, and silicic acid, (b) binders such as, for example, carboxymethylcellulose, gums (*e.g.* alginates, acacia) gelatin, polyvinylpyrrolidone, and sucrose, (c) humectants such as glycerol, (d) disintegrating agents such as agar-agar, calcium carbonate, potato or tapioca starch, alginic acid, certain silicates, and sodium carbonate, (e) solution retarding agents such a paraffin, (f) absorption accelerators such as quaternary ammonium compounds, (g) wetting agents such as, for example, cetyl alcohol and glycerol monostearate, (h) absorbents such as kaolin and bentonite clay, and (i) lubricants such as talc, calcium stearate, magnesium stearate, solid polyethylene glycols, sodium lauryl sulfate, and mixtures thereof. In the case of capsules, tablets and pills, the dosage form may also comprise buffering agents.

Solid compositions of a similar type may also be employed as fillers in soft and hard-filled gelatin capsules using such excipients as lactose or milk sugar as well as high molecular weight polyethylene glycols and the like.

The solid dosage forms of tablets, dragees, capsules, pills, and granules can be prepared with coatings and shells such as enteric coatings and other coatings well known in the pharmaceutical formulating art. They may optionally contain opacifying agents and can also be of a composition that they release the active ingredient(s) only, or preferentially, in a part of the intestinal tract, optionally, in a delayed manner. Exemplary materials include polymers having pH sensitive solubility, including commercially available materials such as Eudragit®. Examples

of embedding compositions which can be used include polymeric substances and waxes.

The active compounds can also be in micro-encapsulated form if appropriate, with one or more of the above-mentioned excipients.

Liquid dosage forms for oral administration include pharmaceutically-acceptable emulsions, solutions, suspensions, syrups and elixirs. In addition to the active compounds, the liquid dosage forms may contain inert diluents commonly used in the art such as, for example, water or other solvents, solubilizing agents and emulsifiers such as ethyl alcohol, isopropyl alcohol, ethyl carbonate, ethyl acetate, benzyl alcohol, benzyl benzoate, propylene glycol, 1,3-butylene glycol, dimethyl formamide, oils (in particular, cottonseed, groundnut, corn, germ, olive, castor, and sesame oils), glycerol, tetrahydrofurfuryl alcohol, polyethylene glycols and fatty acid esters of sorbitan, and mixtures thereof.

Besides inert diluents, the oral compositions can also include adjuvants such as wetting agents, emulsifying and suspending agents, sweetening, flavoring, and perfuming agents.

Suspensions, in addition to the active compounds, may contain suspending agents such as, for example, ethoxylated isostearyl alcohols, polyoxyethylene sorbitol and sorbitan esters, microcrystalline cellulose, aluminum metahydroxide, bentonite, agar-agar, and tragacanth, and mixtures thereof.

Compositions for rectal or vaginal administration are preferably suppositories which can be prepared by mixing the compounds of the present invention with suitable non-irritating excipients or carriers such as cocoa butter, polyethylene glycol or suppository wax, which are solid at room temperature but liquid at body temperature. Accordingly, such carriers melt in the rectum or vaginal cavity, releasing the active compound.

Compounds of the present invention can also be administered in the form of liposomes. As is known in the art, liposomes are generally derived from phospholipids or other lipid substances. Liposomes are formed by mono- or multi-lamellar hydrated liquid crystals that are dispersed in an aqueous medium. Any non-toxic, physiologically-acceptable and metabolizable lipid capable of forming

liposomes can be used. The present compositions in liposome form can contain, in addition to a compound of the present invention, stabilizers, preservatives, excipients, and the like. The preferred lipids are the phospholipids and the phosphatidyl cholines (lecithins), both natural and synthetic. Methods to form liposomes are known in the art. See, for example, Prescott, Ed., Methods in Cell Biology, Volume XIV, Academic Press, New York, N.Y. (1976), p. 33 et seq.

The compounds of the present invention may be used in the form of pharmaceutically-acceptable salts derived from inorganic or organic acids. "Pharmaceutically-acceptable salts" include those salts which are, within the scope of sound medical judgment, suitable for use in contact with the tissues of humans and lower animals without undue toxicity, irritation, allergic response and the like, and are commensurate with a reasonable benefit/risk ratio. Pharmaceutically-acceptable salts are well known in the art. For example, S. M. Berge, *et al.*, describe pharmaceutically-acceptable salts in detail in *J. Pharmaceutical Sciences*, *66*:1 (1977), incorporated herein by reference in its entirety. The salts may be prepared *in situ* during the final isolation and purification of the compounds of the present invention or separately by reacting a free base function with a suitable acid. Representative acid addition salts include, but are not limited to acetate, adipate, alginate, citrate, aspartate, benzoate, benzenesulfonate, bisulfate, butyrate, camphorate, camphorolsulfonate, digluconate, glycerophosphate, hemisulfate, heptanoate, hexanoate, fumarate hydrochloride, hydrobromide, hydroiodide, 2-hydroxyethanesulfonate (isothionate), lactate, maleate, methanesulfonate, nicotinate, 2-naphthalenesulfonate, oxalate, palmoate, pectinate, persulfate, 3-phenylpropionate, picrate, pivalate, propionate, succinate, tartrate, thiocyanate, phosphate, glutamate, bicarbonate, p-toluenesulfonate and undecanoate. Examples of acids which may be employed to form pharmaceutically acceptable acid addition salts include inorganic acids as hydrochloric acid, hydrobromic acid, sulphuric acid and phosphoric acid and such organic acids as oxalic acid, maleic acid, succinic acid and citric acid.

Basic nitrogen-containing groups can be quaternized with agents such as, for example, lower alkyl halides including methyl, ethyl, propyl, and butyl chlorides, bromides and iodides; dialkyl sulfates like dimethyl, diethyl, dibutyl and

diamyl sulfates; long chain halides such as decyl, lauryl, myristyl and stearyl chlorides, bromides and iodides; arylalkyl halides like benzyl and phenethyl bromides and others. Water or oil-soluble or dispersible products are thereby obtained.

Basic addition salts can be prepared *in situ* during the final isolation and purification of compounds of the present invention by reacting a carboxylic acid-containing moiety with a suitable base such as the hydroxide, carbonate or bicarbonate of a pharmaceutically acceptable metal cation or with ammonia or with an organic primary, secondary or tertiary amine. Pharmaceutically-acceptable basic addition salts include, but are not limited to, cations based on alkali metals or alkaline earth metals such as lithium, sodium, potassium, calcium, magnesium and aluminum salts and the like and nontoxic quaternary ammonia and amine cations including ammonium, tetramethylammonium, tetraethylammonium, methylamine, dimethylamine, trimethylamine, triethylamine, diethylamine, ethylamine and the like. Other representative organic amines useful for the formation of base addition salts include ethylenediamine, ethanolamine, diethanolamine, piperidine, piperazine and the like.

Dosage forms for topical administration of a compound of the present invention include powders, sprays, ointments and inhalants. The active compound is mixed under sterile conditions with a pharmaceutically-acceptable carrier and any needed preservatives, buffers, or propellants which may be required. Ophthalmic formulations, eye ointments, powders, and solutions are also contemplated as being within the scope of the present invention.

Actual dosage levels of active ingredients in the pharmaceutical compositions of this present invention may be varied so as to obtain an amount of the active compound(s) that is effective to achieve the desired therapeutic response for a particular patient, compositions, and mode of administration. The selected dosage level will depend upon the activity of the particular compound, the route of administration, the severity of the condition being treated, and the condition and prior medical history of the patient being treated. However, it is within the skill of the art to start doses of the compound at levels lower than required to achieve the desired therapeutic effort and to gradually increase the dosage until the desired effect is achieved.

Generally dosage levels of about 0.1 to about 1000 mg, about 0.5 to about 500 mg, about 1 to about 250 mg, about 1.5 to about 100mg, and preferably of about 5 to about 20 mg of active compound per kilogram of body weight per day are administered orally or intravenously to a mammalian patient. If desired, the effective daily dose may be divided into multiple doses for purposes of administration, *e.g.*, two to four separate doses per day.

The efficacy of the compounds of the present invention have been investigated and can be described by parameters, such as, for example EC50 and LC50. As used herein, the term EC50 refers to the effective concentration needed to inhibit activity by 50% in a cell based assay. The term IC50, as used herein, refers to the concentration required to inhibit protein activity in a biochemical assay by 50%. The term LD50, as used herein, refers to the compound concentration necessary to kill 50% of the cells over a defined time interval in toxicity assays.

TABLE 1

Proteins which Comprise I or A domains, G proteins, heterotrimeric G proteins, and tubulin GTPase.

1. TIM beta/alpha-barrel (23)

contains parallel beta-sheet barrel, closed; n=8, S=8; strand order 12345678 the first six superfamilies have similar phosphate-binding sites

   1. Triosephosphate isomerase (TIM) (1)
      1. Triosephosphate isomerase (TIM) (12)
   2. Ribulose-phoshate binding barrel (4)
      1. Histidine biosynthesis enzymes (2)

         structural evidence for the gene duplication within the barrel fold
      2. D-ribulose-5-phosphate 3-epimerase (1)
      3. Orotidine 5'-monophosphate decarboxylase (OMP decarboxylase) (4)
      4. Tryptophan biosynthesis enzymes (6)
   3. Thiamin phosphate synthase (1)

TABLE 1 (continued)

1. Thiamin phosphate synthase (1)

4. FMN-linked oxidoreductases (1)

1. FMN-linked oxidoreductases (9)

5. Inosine monophosphate dehydrogenase (IMPDH) (1)

The phosphape moiety of substrate binds in the 'common' phosphate-binding site

1. Inosine monophosphate dehydrogenase (IMPDH) (4)

6. PLP-binding barrel (2)

circular permutation of the canonical fold: begins with an alpha helix and ends with a beta-strand

1. Alanine racemase-like, N-terminal domain (4)
2. "Hypothetical" protein ybl036c (1)

7. NAD(P)-linked oxidoreductase (1)

1. Aldo-keto reductases (NADP) (7)

Common fold covers whole protein structure

8. (Trans)glycosidases (7)

1. alpha-Amylases, N-terminal domain (22)

Common fold domain is interrupted by a small calcium-binding subdomain

This domain is followed by an all-beta domain common to the family

2. beta-Amylase (4)
3. beta-glycanases (21)

consist of a number of sequence families

4. Family 1 of glycosyl hydrolase (8)
5. Type II chitinase (9)

glycosylase family 18

6. Bacterial chitobiase (beta-N-acetylhexosaminidase), catalytic domain (1)

Glycosyl hydrolase family 20

TABLE 1 (continued)

7. Beta-D-glucan exohydrolase, N-terminal domain (1)

9. Metallo-dependent hydrolases (3)

the beta-sheet barrel is similarly distorted and capped by a C-terminal helix has transition metal ions bound inside the barrel

1. Adenosine deaminase (ADA) (1)
2. alpha-subunit of urease, catalytic domain (2)
3. Phosphotriesterase-like (2)

10. Aldolase (4)

Common fold covers whole protein structure

1. Class I aldolase (14)

   the catalytic lysine forms schiff-base intermediate with substrate
2. Class II aldolase (1)

   metal-dependent
3. 5-aminolaevulinate dehydratase, ALAD (porphobilinogen synthase) (3)

   hybrid of classes I and II aldolase
4. Class I DAHP synthetase (2)

11. Enolase C-terminal domain-like (2)

binds metal ion (magnesium or manganese) in conserved site inside barrel

N-terminal alpha+beta domain is common to this family

1. Enolase (2)
2. D-glucarate dehydratase-like (6)

12. Phosphoenolpyruvate/pyruvate domain (6)

1. Pyruvate kinase (5)
2. Pyruvate phosphate dikinase, C-terminal domain (1)
3. Phosphoenolpyruvate carboxylase (1)
4. Phosphoenolpyruvate mutase (1)

   forms a swapped dimer

TABLE 1 (continued)

5. 2-dehydro-3-deoxy-galactarate aldolase (1)

forms a swapped dimer; contains a PK-type metal-binding site

6. Isocitrate lyase (2)

forms a swapped dimer; elaborated with additional subdomains

13. Malate synthase G (1)

1. Malate synthase G (1)

14. RuBisCo, C-terminal domain (1)

1. RuBisCo, large subunit, C-terminal domain (6)

N-terminal domain is alpha+beta

15. Xylose isomerase-like (3)

different families share similar but non-identical metal-binding sites

1. Endonuclease IV (1)
2. L-rhamnose isomerase (1)
3. Xylose isomerase (12)

16. Bacterial luciferase-like (3)

consists of clearly related families of somewhat different folds

1. Bacterial luciferase (alkanal monooxygenase) (1)

typical (beta/alpha)8-barrel fold

2. Non-fluorescent flavoprotein (luxF, FP390) (2)

incomplete beta/alpha barrel with mixed beta-sheet of 7 strands

3. Coenzyme F420 dependent tetrahydromethanopterin reductase (1)

17. Quinolinic acid phosphoribosyltransferase, C-terminal domain (1)

incomplete beta/alpha barrel with parallel beta-sheet of 7 strands

1. Quinolinic acid phosphoribosyltransferase, C-terminal domain (2)

18. Phosphatidylinositol-specific phospholipase C (PI-PLC) (2)

1. Mammalian PLC (1)

TABLE 1 (continued)

2. Bacterial PLC (2)

19. Cobalamin (vitamin B12)-dependent enzymes (3)

1. Methylmalonyl-CoA mutase, N-terminal (CoA-binding) domain (1)

2. Glutamate mutase, large subunit (1)

3. Diol dehydratase, alpha subunit (1)

20. tRNA-guanine transglycosylase (1)

1. tRNA-guanine transglycosylase (1)

21. Dihydropteroate synthetase-like (2)

1. Dihydropteroate synthetase (3)

2. Methyltetrahydrofolate: corrinoid/iron-sulfur protein methyltransferase MetR (1)

22. Uroporphyrinogen decarboxylase, UROD (1)

1. Uroporphyrinogen decarboxylase, UROD (1)

23. Methylenetetrahydrofolate reductase (1)

1. Methylenetetrahydrofolate reductase (1)

2. NAD(P)-binding Rossmann-fold domains (1)

core: 3 layers, a/b/a; parallel beta-sheet of 6 strands, order 321456

The nucleotide-binding modes of this and the next two folds/superfamilies are similar

1. NAD(P)-binding Rossmann-fold domains (8)

1. Alcohol/glucose dehydrogenases, C-terminal domain (9)

N-terminal all-beta domain defines family

2. Tyrosine-dependent oxidoreductases (27)

also known as short-chain dehydrogenases and SDR family

parallel beta-sheet is extended by 7th strand, order 3214567; left-handed

crossover connection between strands 6 and 7

3. Glyceraldehyde-3-phosphate dehydrogenase-like, N-terminal domain (20)

TABLE 1 (continued)

family members also share a common alpha+beta fold in C-terminal domain

4. Formate/glycerate dehydrogenases, NAD-domain (9)

this domain interrupts the other domain which defines family

5. Lactate & malate dehydrogenases, N-terminal domain (16)

6. 6-phosphogluconate dehydrogenase-like, N-terminal domain (8)

the beta-sheet is extended to 8 strands, order 32145678; strands 7 & 8 are antiparallel to the rest

C-terminal domains also show some similarity

7. Amino-acid dehydrogenase-like, C-terminal domain (11)

8. Succinyl-CoA synthetase, alpha-chain, N-terminal (CoA-binding) domain (2)

3. FAD/NAD(P)-binding domain (1)

core: 3 layers, b/b/a; central parallel beta-sheet of 5 strands, order 32145; top antiparallel beta-sheet of 3 strands, meander

1. FAD/NAD(P)-binding domain (5)

1. C-terminal domain of adrenodoxin reductase-like (3)

2. FAD-linked reductases, N-terminal domain (10)

C-terminal domain is alpha+beta is common for the family

3. Guanine nucleotide dissociation inhibitor, GDI (1)

Similar to FAD-linked reductases in both domains but does not bind FAD

4. Succinate dehydrogenase/fumarate reductase N-terminal domain (5)

5. FAD/NAD-linked reductases, N-terminal and central domains (17)

duplication: both domains have similar folds and functions most members of the family contain common C-terminal alpha+beta domain

TABLE 1 (continued)

4. Nucleotide-binding domain (1)
3 layers: a/b/a; parallel beta-sheet of 5 strands, order 32145; Rossmann-like
1. Nucleotide-binding domain (2)
this superfamily shares the common nucleotide-binding site with and provides a link between the Rossmann-fold NAD(P)-binding and FAD/NAD(P)-binding domains
1. N-terminal domain of adrenodoxin reductase-like (3)
2. D-amino acid oxidase, N-terminal domain (2)
This family is probably related to the FAD-linked reductases and shares with them the C-terminal domain fold

5. N-terminal domain of MurD (UDP-N-acetylmuramoyl-L-alanine:D-glutamate ligase) (1)
3 layers: a/b/a; parallel beta-sheet of 5 strands, order 32145; incomplete Rossmann-like fold; binds UDP group
1. N-terminal domain of MurD (UDP-N-acetylmuramoyl-L-alanine:D-glutamate ligase) (1)
1. N-terminal domain of MurD (UDP-N-acetylmuramoyl-L-alanine:D-glutamate ligase) (1)

6. Cellulases (1)
variant of beta/alpha barrel; parallel beta-sheet barrel, closed, n=7, S=8; strand order 1234567
1. Cellulases (1)
1. Cellulases (4)

7. PFL-like glycyl radical enzymes (1)
contains: barrel, closed; n=10, S=10; accommodates a hairpin loop inside the barrel
1. PFL-like glycyl radical enzymes (3)
duplication: the N- and C-terminal halves have similar topologies
1. Pyruvate formate-lyase, PFL (1)
2. R1 subunit of ribonucleotide reductase, C-terminal domain (1)

TABLE 1 (continued)

3. Class III anaerobic ribonucleotide triphosphate reductase NRDD subunit (1)

8. The "swivelling" beta/beta/alpha domain (5)

3 layers: b/b/a; the central sheet is parallel, and the other one is antiparallel; there are some variations in topology

this domain is thought to be mobile in all proteins known to contain it

1. Phosphohistidine domain (2)

contains barrel, closed, n=7, S=10

1. Pyruvate phosphate dikinase, central domain (1)
2. N-terminal domain of enzyme I of the PEP:sugar phosphotransferase system (1)

2. Aconitase, C-terminal domain (1)

contains mixed beta-sheet barrel, closed n=7, S=10

1. Aconitase, C-terminal domain (2)

3. Carbamoyl phosphate synthetase, small subunit N-terminal domain (1)

1. Carbamoyl phosphate synthetase, small subunit N-terminal domain (1)

4. Transferrin receptor ectodomain, apical domain (1)

1. Transferrin receptor ectodomain, apical domain (1)

5. GroEL-like chaperone, apical domain (2)

1. GroEL (2)
2. Group II chaperonin (CCT, TRIC) (1)

9. Barstar-like (2)

2 layers, a/b; parallel beta-sheet of 3 strands, order 123

1. Barstar (barnase inhibitor) (1)

1. Barstar (barnase inhibitor) (1)

2. Ribosomal protein L32e (1)

1. Ribosomal protein L32e (1)

contains irregular N-terminal extension to the common fold

TABLE 1 (continued)

10. Leucine-rich repeat, LRR (right-handed beta-alpha superhelix) (2)
2 curved layers, a/b; parallel beta-sheet: order 1234...N
1. RNI-like (3)
regular structure consisting of similar repeats
1. Ribonuclease inhibitor (2)
2. Rna1p (1)
3. Cyclin A/CDK2-associated p19, Skp2 (1)
2. L domain-like (5)
less regular structure consisting of variable repeats
1. Internalin B LRR domain (1)
2. Rab geranylgeranyltransferase alpha-subunit, C-terminal domain (1)
3. mRNA export factor tap (1)
4. U2A'-like (1)
duplication: consists of 5-6 partly irregular repeats
5. L1 and L2 domains of the type 1 insulin-like growth factor receptor (1)

11. Outer arm dynein light chain 1 (1)
(beta-beta-alpha)n superhelix
1. Outer arm dynein light chain 1 (1)
1. Outer arm dynein light chain 1 (1)

12. Ribosomal proteins L15p and L18e (1)
core: three turns of irregular (beta-beta-alpha)n superhelix
1. Ribosomal proteins L15p and L18e (1)
1. Ribosomal proteins L15p and L18e (2)

13. SpoIIaa-like (2)
core: 4 turns of a (beta-alpha)n superhelix
1. C-terminal domain of phosphatidylinositol transfer protein sec14p (1)
1. C-terminal domain of phosphatidylinositol transfer protein sec14p (1)

TABLE 1 (continued)

2. SpoIIaa (1)

1. SpoIIaa (1)

14. ClpP/crotonase (1)

core: 4 turns of (beta-beta-alpha)n superhelix

1. ClpP/crotonase (3)

1. Clp protease, ClpP subunit (1)
2. Photosystem II D1 C-terminal processing protease, catalytic domain (1)
3. Crotonase-like (4)

15. BRCT domain (1)

3 layers, a/b/a; core: parallel beta-sheet of 4 strands, order 2134

1. BRCT domain (2)

1. DNA-repair protein XRCC1 (1)
2. NAD+-dependent DNA ligase, domain 4 (1)

16. beta-subunit of the lumazine synthase/riboflavin synthase complex (1)

3 layers, a/b/a; core: parallel beta-sheet of 4 strands, order 2134

1. beta-subunit of the lumazine synthase/riboflavin synthase complex (1)

1. beta-subunit of the lumazine synthase/riboflavin synthase complex (4)

17. Caspase-like (1)

3 layers, a/b/a; core: parallel beta-sheet of 4 strands, order 2134

1. Caspase-like (2)

heterodimeric protein folded in a single domain

1. Caspase (3)
2. Gingipain R (RgpB), N-terminal domain (1)

18. DNA glycosylase (1)

3 layers, a/b/a; core: parallel beta-sheet of 4 strands, order 2134

1. DNA glycosylase (2)

1. Uracil-DNA glycosylase (3)
2. G:T/U mismatch-specific DNA glycosylase (1)

TABLE 1 (continued)

19. Catalytic domain of malonyl-CoA ACP transacylase (1)
3 layers, a/b/a; core: parallel beta-sheet of 4 strands, order 2134
1. Catalytic domain of malonyl-CoA ACP transacylase (1)
1. Catalytic domain of malonyl-CoA ACP transacylase (1)

20. Initiation factor IF2/eIF5b, domain 3 (1)
3 layers, a/b/a; core: parallel beta-sheet of 4 strands, order 2134
1. Initiation factor IF2/eIF5b, domain 3 (1)
1. Initiation factor IF2/eIF5b, domain 3 (1)

21. Ribosomal protein L13 (1)
3 layers, a/b/a; core: parallel beta-sheet of 4 strands, order 3214
1. Ribosomal protein L13 (1)
1. Ribosomal protein L13 (1)

22. Ribosomal protein L4 (1)
3 layers, a/b/a; core: parallel beta-sheet of 4 strands, order 1423
1. Ribosomal protein L4 (1)
1. Ribosomal protein L4 (2)

23. Flavodoxin-like (16)
3 layers, a/b/a; parallel beta-sheet of 5 strand, order 21345
1. CheY-like (3)
1. CheY-related (11)
2. Receiver domain of the ethylene receptor (1)
3. Negative regulator of the amidase operon AmiR (1)
2. Toll/Interleukin receptor TIR domain (1)
1. Toll/Interleukin receptor TIR domain (2)
3. Hypothetical protein MTH538 (1)
1. Hypothetical protein MTH538 (1)
4. Succinyl-CoA synthetase domains (1)
1. Succinyl-CoA synthetase domains (4)
contain additional N-terminal strand "0", antiparallel to strand 2

TABLE 1 (continued)

5. Flavoproteins (3)
    1. Flavodoxin-related (8)
    binds FMN
    2. NADPH-cytochrome p450 reductase, N-terminal domain (2)
    3. Quinone reductase (4)
    binds FAD
6. Cobalamin (vitamin B12)-binding domain (1)
    1. Cobalamin (vitamin B12)-binding domain (4)
7. Ornithine decarboxylase N-terminal "wing" domain (1)
    1. Ornithine decarboxylase N-terminal "wing" domain (1)
8. N5-carboxyaminoimidazole ribonucleotide (N5-CAIR) mutase PurE (1)
    1. N5-carboxyaminoimidazole ribonucleotide (N5-CAIR) mutase PurE (1)
9. Cutinase-like (1)
    1. Cutinase-like (3)
    this family can be also classified into alpha/beta hydrolase superfamily
10. Esterase/acetylhydrolase (4)
    1. Esterase (1)
    2. Esterase domain of haemagglutinin-esterase-fusion glycoprotein HEF1 (1)
    3. Acetylhydrolase (1)
    4. Rhamnogalacturonan acetylesterase (1)
11. Beta-D-glucan exohydrolase, C-terminal domain (1)
    1. Beta-D-glucan exohydrolase, C-terminal domain (1)
12. Formate/glycerate dehydrogenase catalytic domain-like (3)
    1. Formate/glycerate dehydrogenases, substrate-binding domain (6)
    this domain is interrupted by the Rossmann-fold domain

TABLE 1 (continued)

2. L-alanine dehydrogenase (1)
3. S-adenosylhomocystein hydrolase (2)

13. Type II 3-dehydroquinate dehydratase (1)
1. Type II 3-dehydroquinate dehydratase (2)

14. Nucleoside 2-deoxyribosyltransferase (1)
1. Nucleoside 2-deoxyribosyltransferase (1)

15. Ribosomal protein S2 (1)
fold elaborated with additional structures
1. Ribosomal protein S2 (1)

16. Class I glutamine amidotransferase-like (4)
conserved positions of the oxyanion hole and catalytic nucleophile; different constituent families contain different additional structures
1. Class I glutamine amidotransferases (GAT) (3)
contains a catalytic Cys-His-Glu triad
2. Intracellular protease (1)
contains a catalytic Cys-His-Glu triad that differs from the class I GAT triad
3. Catalase, C-terminal domain (1)
4. Aspartyl dipeptidase PepE (1)
probable circular permutation in the common core; contains a catalytic Ser-His-Glu triad

24. Methylglyoxal synthase-like (1)
3 layers, a/b/a; parallel beta-sheet of 5 strands, order 32145
1. Methylglyoxal synthase-like (2)
contains a common phosphate-binding site
1. Carbamoyl phosphate synthetase, large subunit allosteric, C-terminal domain (1)
2. Methylglyoxal synthase, MgsA (1)

25. Ferredoxin reductase-like, C-terminal NADP-linked domain (1)
3 layers, a/b/a; parallel beta-sheet of 5 strands, order 32145

TABLE 1 (continued)

1. Ferredoxin reductase-like, C-terminal NADP-linked domain (5)
binds NADP differently than classical Rossmann-fold
N-terminal FAD-linked domain contains (6,10) barrel
1. Reductases (10)
2. Phthalate dioxygenase reductase (1)
contains additional 2Fe-2S ferredoxin domain
3. Dihydroorotate dehydrogenase B, PyrK subunit (1)
contains 2Fe-2S cluster in the C-terminal extension
4. NADPH-cytochrome p450 reductase-like (2)
5. Flavohemoglobin, C-terminal domain (1)
contains additional globin domain

26. Adenine nucleotide alpha hydrolase-like (3)
core: 3 layers, a/b/a ; parallel beta-sheet of 5 strands, order 32145
1. Nucleotidylyl transferase (3)
1. Class I aminoacyl-tRNA synthetases (RS), catalytic domain (10)
contains a conserved all-alpha subdomain at the C-terminal extension
2. Cytidylyltransferase (1)
3. Adenylyltransferase (2)
2. Adenine nucleotide alpha hydrolases (2)
1. N-type ATP pyrophosphatases (3)
2. Phosphoadenylyl sulphate (PAPS) reductase (1)
3. UDP-glucose dehydrogenase (UDPGDH), C-terminal (UDP-binding) domain (1)
1. UDP-glucose dehydrogenase (UDPGDH), C-terminal (UDP-binding) domain (1)

27. Pyrimidine nucleoside phosphorylase central domain (1)
3 layers: a/b/a; parallel beta-sheet of 5 strands, order 32145; Rossmann-like
1. Pyrimidine nucleoside phosphorylase central domain (1)

TABLE 1 (continued)

1. Pyrimidine nucleoside phosphorylase central domain (2)

28. N-terminal domain of DNA photolyase (1)

3 layers: a/b/a; parallel beta-sheet of 5 strands, order 32145; Rossmann-like

1. N-terminal domain of DNA photolyase (1)

1. N-terminal domain of DNA photolyase (2)

29. ETFP adenine nucleotide-binding domain-like (1)

3 layers: a/b/a, core: parallel beta-sheet of 5 strands, order 32145

1. ETFP adenine nucleotide-binding domain-like (2)

1. Electron transfer flavoprotein, ETFP (2)

contains additional strands on both edges of the core sheet

2. "Hypothetical" protein MJ0577 (1)

30. Biotin carboxylase N-terminal domain-like (1)

3 layers: a/b/a; parallel or mixed beta-sheet of 4 to 6 strands

possible rudiment form of Rossmann-fold domain

1. Biotin carboxylase N-terminal domain-like (5)

superfamily defined by the common ATP-binding domain that follows this one

1. Biotin carboxylase/Carbamoyl phosphate synthetase (5)
2. D-Alanine ligase N-terminal domain (2)
3. Prokaryotic glutathione synthetase, N-terminal domain (1)
4. Eukaryotic glutathione synthetase (1)

circularly permuted version of prokaryotic enzyme

5. Synapsin Ia domain (1)

31. DHS-like NAD/FAD-binding domain (1)

3 layers: a/b/a; parallel beta-sheet of 6 strands, order 321456; Rossmann-like

1. DHS-like NAD/FAD-binding domain (4)

binds cofactor molecules in the opposite direction than classical Rossman fold

1. Deoxyhypusine synthase, DHS (1)

TABLE 1 (continued)

2. C-terminal domain of the electron transfer flavoprotein alpha subunit (2)
lacks strand 3; shares the FAD-binding mode with the pyruvate oxidase domain
3. Pyruvate oxidase and decarboxylase, middle domain (5)
N-terminal domain is Pyr module, and C-terminal domain is PP module of thiamin diphosphate-binding fold
4. Transhydrogenase domain III (dIII) (3)
binds NADP, shares with the pyruvate oxidase FAD-binding domain a common ADP-binding mode

32. Tubulin, GTPase domain (1)
3 layers: a/b/a; parallel beta-sheet of 6 strands, order 321456
1. Tubulin, GTPase domain (1)
1. Tubulin, GTPase domain (3)

33. Cysteine hydrolase (1)
3 layers: a/b/a; parallel beta-sheet of 6 strands, order 321456
1. Cysteine hydrolase (2)
1. N-carbamoylsarcosine amidohydrolase (1)
2. YcaC (1)

34. Halotolerance protein Hal3 (1)
3 layers: a/b/a; parallel beta-sheet of 6 strands, order 321456
1. Halotolerance protein Hal3 (1)
1. Halotolerance protein Hal3 (1)

35. Glucosamine 6-phosphate deaminase/isomerase (1)
3 layers: a/b/a; parallel beta-sheet of 6 strands, order 324561
1. Glucosamine 6-phosphate deaminase/isomerase (1)
1. Glucosamine 6-phosphate deaminase/isomerase (2)

36. Thiamin diphosphate-binding fold (THDP-binding) (1)
3 layers: a/b/a; parallel beta-sheet of 6 strands, order 213465
1. Thiamin diphosphate-binding fold (THDP-binding) (4)

TABLE 1 (continued)

both pyridine (Pyr)- and pyrophosphate (PP)-binding modules have this fold

conserved core consists of two Pyr and two PP-modules and binds two coenzyme molecules

1. Pyruvate oxidase and decarboxylase (5)

   Pyr module is N-terminal domain, PP module is C-terminal domain

   Rossmann-like domain is between them

2. Transketolase, TK (1)

3. Branched-chain alpha-keto acid dehydrogenase (2)

   parent family to TK and PFOR

   heterodimeric protein related to TK; alpha-subunit is the PP module and the N-terminal domain of beta-subunit is the Pyr module

4. Pyruvate-ferredoxin oxidoreductase, PFOR, domains I and VI (1)

   domains VI, I and II are arranged in the same way as the TK N, M and C domains

37. P-loop containing nucleotide triphosphate hydrolases (1)

3 layers: a/b/a, parallel or mixed beta-sheets of variable sizes

1. P-loop containing nucleotide triphosphate hydrolases (14)

   division into families based on beta-sheet topologies

   1. Nucleotide and nucleoside kinases (16)

      parallel beta-sheet of 5 strands, order 23145

   2. Shikimate kinase (1)

      similar to the nucleotide/nucleoside kinases but acts on different substrate

   3. Chloramphenicol phosphotransferase (1)

      similar to the nucleotide/nucleoside kinases but acts on different substrate

TABLE 1 (continued)

4. Adenosine-5'phosphosulfate kinase (APS kinase) (1)
5. PAPS sulfotransferase (4)

similar to the nucleotide/nucleoside kinases but transfer sulphate group

6. Phosphoribulokinase/pantothenate kinase (2)
7. 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase, kinase domain (1)
8. G proteins (28)

core: mixed beta-sheet of 6 strands, order 231456; strand 2 is antiparallel to the rest

9. Motor proteins (7)
10. Nitrogenase iron protein-like (10)

core: parallel beta-sheet of 7 strands; order 3241567

11. RecA protein-like (ATPase-domain) (9)

core: mixed beta-sheet of 8 strands, order 32451678; strand 7 is antiparallel to the rest

12. ABC transporter ATPase domain-like (7)

there are two additional subdomains inserted into the central core that has a RecA-like topology

13. Extended AAA-ATPase domain (13)

fold is similar to that of RecA, but lacks the last two strands, followed by a family-specific all-alpha Arg-finger domain

14. RNA helicase (1)

duplication: consists of two similar domains, one binds NTP and the other binds RNA; also contains an all-alpha subdomain in the C-terminal extension

38. Fructose permease, subunit IIb (1)

3 layers: a/b/a, parallel beta-sheet of 6 strands, order 324156

1. Fructose permease, subunit IIb (1)
    1. Fructose permease, subunit IIb (1)

TABLE 1 (continued)

39. Nicotinate mononucleotide:5,6-dimethylbenzimidazole phosphoribosyltransferase (CobT) (1)
3 layers: a/b/a; parallel beta-sheet of 7 strands, order 3214567
1. Nicotinate mononucleotide:5,6-dimethylbenzimidazole phosphoribosyltransferase (CobT) (1)
1. Nicotinate mononucleotide:5,6-dimethylbenzimidazole phosphoribosyltransferase (CobT) (1)

40. Methylesterase CheB, C-terminal domain (1)
3 layers: a/b/a, parallel beta-sheet of 7 strands, order 3421567
1. Methylesterase CheB, C-terminal domain (1)
1. Methylesterase CheB, C-terminal domain (1)

41. Subtilisin-like (1)
3 layers: a/b/a, parallel beta-sheet of 7 strands, order 2314567; left-handed crossover connection between strands 2 & 3
1. Subtilisin-like (2)
1. Subtilases (12)
2. Serine-carboxyl proteinase PSCP (1)
elaborated with additional structures

42. Arginase/deacetylase (1)
3 layers: a/b/a, parallel beta-sheet of 8 strands, order 21387456
1. Arginase/deacetylase (2)
1. Arginase (2)
2. Histone deacetylase, HDAC (1)

43. CoA-dependent acyltransferases (1)
core: 2 layers, a/b; mixed beta-sheet of 6 strands, order 324561; strands 3 & 6 are antiparallel to the rest
1. CoA-dependent acyltransferases (1)
1. CoA-dependent acyltransferases (5)

44. Phosphotyrosine protein phosphatases I-like (2)
3 layers: a/b/a; parallel beta-sheet of 4 strands, order 2134

TABLE 1 (continued)

1. Phosphotyrosine protein phosphatases I (1)
share the common active site structure with the family II
1. Low-molecular-weight phosphotyrosine protein phosphatases (3)

2. Enzyme IIB-cellobiose (1)
1. Enzyme IIB-cellobiose (1)

45. (Phosphotyrosine protein) phosphatases II (1)
core: 3 layers, a/b/a; parallel beta-sheet of 4 strands, order 1432
1. (Phosphotyrosine protein) phosphatases II (3)
share with the family I the common active site structure with a circularly permuted topology
1. Dual-specificity phosphatases (2)
2. Higher-molecular-weight phosphotyrosine protein phosphatases (8)
have an extension to the beta-sheet of 3 antiparallel strands before strand 4
3. Phoshphoinositide phosphatase Pten (Pten tumor suppressor), N-terminal domain (1)

46. Rhodanese/Cell cycle control phosphatase (1)
3 layers: a/b/a; parallel beta-sheet of 5 strands, order 32451
1. Rhodanese/Cell cycle control phosphatase (2)
the active site structure is similar to those of the families I and II protein phosphatases; the topology can be related by a different circular permutation to the family I topology
1. Cell cycle control phosphatase, catalytic domain (2)
2. Sulfurtransferase (rhodanese) (2)
duplication, consists of two domains of this fold

47. Thioredoxin fold (3)
core: 3 layers, a/b/a; mixed beta-sheet of 4 strands, order 4312; strand 3 is antiparallel to the rest

TABLE 1 (continued)

1. Thioredoxin-like (10)
    1. Thioltransferase (12)
    2. PDI-like (3)

    duplication: contains two tandem repeats of this fold
    3. Calsequestrin (1)

    duplication: contains three tandem repeats of this fold
    4. Disulphide-bond formation facilitator (DSBA) (2)
    5. Glutathione S-transferases, N-terminal domain (23)
    6. Phosducin (2)
    7. Endoplasmic reticulum protein ERP29, N-domain (1)
    8. spliceosomal protein U5-15Kd (1)
    9. Disulfide bond isomerase, DsbC, C-terminal domain (1)

    elaborated common fold
    10. Glutathione peroxidase-like (6)
2. RNA 3'-terminal phosphate cyclase, RPTC, insert domain (1)
    1. RNA 3'-terminal phosphate cyclase, RPTC, insert domain (1)
3. Thioredoxin-like 2Fe-2S ferredoxin (1)
    1. Thioredoxin-like 2Fe-2S ferredoxin (1)

48. Transketolase C-terminal domain-like (1)

3 layers: a/b/a; mixed beta-sheet of 5 strands, order 13245, strand 1 is antiparallel to the rest

1. Transketolase C-terminal domain-like (3)
    1. Transketolase (1)
    2. Branched-chain alpha-keto acid dehydrogenase beta-subunit, C-domain (2)
    3. Pyruvate-ferredoxin oxidoreductase, PFOR, domain II (1)

49. Pyruvate kinase C-terminal domain-like (2)

3 layers: a/b/a; mixed beta-sheet of 5 strands, order 32145, strand 5 is antiparallel to the rest

1. Pyruvate kinase, C-terminal domain (1)

TABLE 1 (continued)

1. Pyruvate kinase, C-terminal domain (5)
2. ATP syntase (F1-ATPase), gamma subunit (1)

contains an antiparallel coiled coil formed by – anb C-terminal extensions to the common fold

1. ATP syntase (F1-ATPase), gamma subunit (2)

50. Leucine aminopeptidase, N-terminal domain (1)

3 layers: a/b/a; mixed beta-sheet of 5 strands, order 23145; strand 2 is antiparallel to the rest

1. Leucine aminopeptidase, N-terminal domain (1)
    1. Leucine aminopeptidase, N-terminal domain (1)

51. Anticodon-binding domain-like (4)

3 layers: a/b/a; mixed beta-sheet of five strands, order 21345; strand 4 is antiparallel to the rest

1. Anticodon-binding domain of Class II aaRS (1)
    1. Anticodon-binding domain of Class II aaRS (5)
2. TolB, N-terminal domain (1)
    1. TolB, N-terminal domain (1)
3. Diol dehydratase, beta subunit (1)
    1. Diol dehydratase, beta subunit (1)

    contains additional structures in the C-terminal extension
4. Maf/Ham1 (2)

    elaborated with additional structures inserted in the common fold
    1. Ham1 (1)
    2. Maf protein (1)

52. Restriction endonuclease-like (3)

core: 3 layers, a/b/a; mixed beta-sheet of 5 strands, order 12345; strands 2 &, in some families, 5 are antiparallel to the rest

1. Restriction endonuclease-like (17)
    1. Restriction endonuclease EcoRI (1)
    2. Restriction endonuclease EcoRV (1)

TABLE 1 (continued)

3. Restriction endonuclease BamHI (1)
4. Restriction endonuclease BglI (1)
5. Restriction endonuclease BglII (1)
6. Restriction endonuclease PvuII (1)
7. Restriction endonuclease Cfr10I (1)
8. Restriction endonuclease MunI (1)
9. Restriction endonuclease NaeI (1)
10. Restriction endonuclease NgoIV (1)
11. Restriction endonuclease BsobI (1)
12. Restriction endonuclease FokI, C-terminal (catalytic) domain (1)
13. lambda exonuclease (1)
14. DNA mismatch repair protein MutH from (1)
15. Very short patch repair (VSR) endonuclease (1)
16. TnsA endonuclease, N-terminal domain (1)
17. Holliday junction resolvase (Endonuclease I) (1)

2. tRNA splicing endonuclease, C-terminal domain (1)
   1. tRNA splicing endonuclease, C-terminal domain (1)
3. Eukaryotic RPB5 N-terminal domain (1)
   1. Eukaryotic RPB5 N-terminal domain (1)

53. Resolvase-like (2)

Core: 3 layers: a/b/a; mixed beta-sheet of 5 strands, order 21345; strand 5 is antiparallel to the rest

1. Resolvase-like (2)
   1. gamma, delta resolvase, large fragment (1)
   2. 5' to 3' exonuclease (5)

      contains additional strand and alpha-helical arch; strand order 321456; strand 6 is antiparallel to the rest
2. beta-carbonic anhydrase (1)
   1. beta-carbonic anhydrase (2)

TABLE 1 (continued)

54. IIA domain of mannose transporter, IIA-Man (1)

3 layers: a/b/a; mixed beta-sheet of 5 strands, order 21345; strand 5 is antiparallel to the rest

1. IIA domain of mannose transporter, IIA-Man (1)

active dimer is formed by strand 5 swapping

1. IIA domain of mannose transporter, IIA-Man (1)

55. Ribonuclease H-like motif (7)

3 layers: a/b/a; mixed beta-sheet of 5 strands, order 32145; strand 2 is antiparallel to the rest

1. Actin-like ATPase domain (4)

duplication contains two domains of this fold

1. Actin/HSP70 (8)
2. Acetate kinase (1)
3. Hexokinase (3)
4. Glycerol kinase (1)

2. Creatinase/prolidase N-terminal domain (1)

1. Creatinase/prolidase N-terminal domain (2)

3. Ribonuclease H-like (6)

consists of one domain of this fold

1. Ribonuclease H (4)
2. Retroviral integrase, catalytic domain (3)
3. mu transposase, core domain (1)
4. Transposase inhibitor (Tn5 transposase) (1)
5. DnaQ-like 3'-5' exonuclease (11)
6. RuvC resolvase (1)

4. Translational machinery components (2)

1. Ribosomal protein L18 and S11 (2)
2. Middle domain of eukaryotic peptide chain release factor subunit 1, ERF1 (1)

5. Hypothetical protein MTH1175 (1)

TABLE 1 (continued)

1. Hypothetical protein MTH1175 (1)

6. DNA repair protein MutS, domain II (1)

1. DNA repair protein MutS, domain II (2)

7. Methylated DNA-protein cysteine methyltransferase domain (1)

1. Methylated DNA-protein cysteine methyltransferase domain (3)

56. Phosphorylase/hydrolase-like (6)

core: 3 layers, a/b/a ; mixed sheet of 5 strands: order 21354; strand 4 is antiparallel to the rest; contains crossover loops

1. Hydrogenase maturating endopeptidase HybD (1)

the fold coincides with the consensus core structure

1. Hydrogenase maturating endopeptidase HybD (1)

2. Purine and uridine phosphorylases (1)

complex architecture; contains mixed beta-sheet of 8 strands, order 23415867, strands 3, 6 & 7 are antiparallel to the rest; and barrel, closed; n=5, S=8

1. Purine and uridine phosphorylases (6)

3. Peptidyl-tRNA hydrolase (1)

1. Peptidyl-tRNA hydrolase (1)

4. Pyrrolidone carboxyl peptidase (pyroglutamate aminopeptidase) (1)

1. Pyrrolidone carboxyl peptidase (pyroglutamate aminopeptidase) (2)

5. Zn-dependent exopeptidases (5)

core: mixed beta-sheet of 8 strands, order 12435867; strands 2, 6 & 7 are antiparallel to the rest

1. Pancreatic carboxypeptidases (6)

2. Carboxypeptidase T (1)

3. Leucine aminopeptidase, C-terminal domain (1)

4. Bacterial exopeptidases (3)

5. Transferrin receptor ectodomain, protease-like domain (1)

6. LigB subunit of an aromatic-ring-opening dioxygenase LigAB (1)

TABLE 1 (continued)

circular permutation of the common fold, most similar to the PNP fold

1. LigB subunit of an aromatic-ring-opening dioxygenase LigAB (1)

57. Molybdenumm cofactor biosynthesis protein MogA (1)

3 layers: a/b/a; mixed beta-sheet of 5 strands; order: 21354, strand 5 is antiparallel to the rest; permutation of the Phosphorylase/hydrolase-like fold

1. Molybdenumm cofactor biosynthesis protein MogA (1)

1. Molybdenumm cofactor biosynthesis protein MogA (1)

58. Amino acid dehydrogenase-like, N-terminal domain (1)

3 layers: a/b/a; mixed beta-sheet of 5 strands; 12435, strand 2 is antiparallel to the rest

1. Amino acid dehydrogenase-like, N-terminal domain (3)

1. Amino acid dehydrogenases (7)

dimerisation domain

2. Tetrahydrofolate dehydrogenase/cyclohydrolase (3)

3. Mitochondrial NAD(P)-dependent malic enzyme (1)

this domain is decorated with additional structures; includes N-terminal additional subdomains

59. Glutamate ligase domain (1)

3 layers: a/b/a; mixed beta-sheet of 6 strands, order 126345; strand 1 is antiparallel to the rest

1. Glutamate ligase domain (2)

1. MurD/MurF C-terminal domain (2)

2. Folylpolyglutamate synthetase, C-terminal domain (1)

60. Phosphoglycerate mutase-like (1)

core: 3 layers, a/b/a; mixed beta-sheet of 6 strands, order 324156; strand 5 is antiparallel to the rest

1. Phosphoglycerate mutase-like (4)

1. Phosphoglycerate mutase (1)

2. Acid phosphatase (2)

TABLE 1 (continued)

3. Phytase (myo-inositol-hexakisphosphate-3-phosphohydrolase) (3)
4. 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase, phosphatase domain (1)

61. PRTase-like (1)

core: 3 layers, a/b/a; mixed beta-sheet of 6 strands, order 321456; strand 3 is antiparallel to the rest

1. PRTase-like (2)
    1. Phosphoribosyltransferases (PRTases) (14)
    2. Phosphoribosylpyrophosphate synthetase (1)

    duplication: consists of two domains of this fold

62. Integrin A (or I) domain (1)

core: 3 layers, a/b/a; mixed beta-sheet of 6 strands, order 321456; strand 3 is antiparallel to the rest

1. Integrin A (or I) domain (1)
    1. Integrin A (or I) domain (7)

63. Glutaconate-CoA transferase subunits (1)

core: 3 layers: a/b/a; parallel or mixed b-sheet of 6 strands, order 432156; part of sheet is folded upon itself and forms a barrel-like structure

1. Glutaconate-CoA transferase subunits (1)
    1. Glutaconate-CoA transferase subunits (2)

64. Pyruvate-ferredoxin oxidoreductase, PFOR, domain III (1)

3 layers: a/b/a; mixed beta-sheet of 6 strands, order 231456; strand 3 is antiparallel to the rest

1. Pyruvate-ferredoxin oxidoreductase, PFOR, domain III (1)
    1. Pyruvate-ferredoxin oxidoreductase, PFOR, domain III (1)

65. Formyltransferase (1)

3 layers: a/b/a; mixed beta-sheet of 7 strands, order 3214567; strand 6 is antiparallel to the rest

1. Formyltransferase (1)

TABLE 1 (continued)

1. Formyltransferase (2)

66. S-adenosyl-L-methionine-dependent methyltransferases (1)

core: 3 layers, a/b/a; mixed beta-sheet of 7 strands, order 3214576; strand 7 is antiparallel to the rest

1. S-adenosyl-L-methionine-dependent methyltransferases (11)
    1. Catechol O-methyltransferase, COMT (1)
    2. RNA methyltransferase FtsJ (1)
    3. Fibrillarin homologue (1)
    4. Hypothetical protein MJ0882 (1)
    5. Glycine N-methyltransferase (1)
    6. Arginine methyltransferase, HMT1 (1)

       lacks the last two strands of the common fold replaced with a beta-sandwich oligomerisation subdomain
    7. Protein-L-isoaspartate O-methyltransferase (1)

       another C-terminal variation of the common fold with additional alpha+beta subdomain
    8. Chemotaxis receptor methyltransferase CheR, C-terminal domain (1)

       contains additional N-terminal all-alpha domain, res. 11-91
    9. RNA methylases (3)
    10. DNA methylases (5)
    11. Type II DNA methylase (2)

        circularly permuted version of the common fold

67. PLP-dependent transferases (1)

main domain: 3 layers: a/b/a, mixed beta-sheet of 7 strands, order 3245671; strand 7 is antiparallel to the rest

1. PLP-dependent transferases (5)
    1. AAT-like (9)
    2. Beta-eliminating lyases (2)
    3. Cystathionine synthase-like (8)

TABLE 1 (continued)

4. omega-Amino acid:pyruvate aminotransferase-like (15)
5. Ornithine decarboxylase major domain (1)

68. Nucleotide-diphospho-sugar transferases (1)

3 layers: a/b/a; mixed beta-sheet of 7 strands, order 3214657; strand 6 is antiparallel to the rest

1. Nucleotide-diphospho-sugar transferases (8)
    1. Spore coat polysaccharide biosynthesis protein SpsA (1)
    2. beta 1,4 galactosyltransferase (b4GalT1) (1)
    3. CMP acylneuraminate synthetase (1)
    4. Galactosyltransferase LgtC (1)
    5. N-acetylglucosamine 1-phosphate uridyltransferase GlmU, N-terminal domain (1)
    6. glucose-1-phosphate thymidylyltransferase RmlA (1)
    7. 1,3-Glucuronyltransferase I (glcAT-I) (1)
    8. Molybdenum cofactor biosynthesis protein MobA (1)

69. alpha/beta-Hydrolases (1)

core: 3 layers, a/b/a; mixed beta-sheet of 8 strands, order 12435678, strand 2 is antiparallel to the rest

1. alpha/beta-Hydrolases (20)

many members have left-handed crossover connection between strand 8 and additional strand 9

    1. Acetylcholinesterase-like (8)
    2. Carboxylesterase (2)
    3. Mycobacterial antigens (2)
    4. Prolyl oligopeptidase, C-terminal domain (1)
    5. Serine carboxypeptidase (4)
    6. Gastric lipase (1)
    7. Proline iminopeptidase (2)
    8. Haloalkane dehalogenase (3)
    9. Dienelactone hydrolase (2)

TABLE 1 (continued)

10. Carbon-carbon bond hydrolase (1)
11. Epoxide hydrolase (3)
12. Haloperoxidase (5)
13. Thioesterases (2)
14. Carboxylesterase/thioesterase 1 (2)
15. A novel bacterial esterase (1)
16. Lipase (1)
17. Fungal lipases (9)
18. Bacterial lipase (5)
19. Pancreatic lipase, N-terminal domain (6)
20. Hydroxynitrile lyase (2)

70. Nucleoside hydrolase (1)

core: 3 layers, a/b/a ; mixed beta-sheet of 8 strands, order 32145687; strand 7 is antiparallel to the rest

1. Nucleoside hydrolase (1)
    1. Nucleoside hydrolase (2)

71. Dihydrofolate reductases (1)

3 layers: a/b/a; mixed beta-sheet of 8 strands, order 34251687; strand 8 is antiparallel to the rest

1. Dihydrofolate reductases (1)
    1. Dihydrofolate reductases (10)

72. Ribokinase-like (2)

core: 3 layers: a/b/a; mixed beta-sheet of 8 strands, order 21345678, strand 7 is antiparallel to the rest

potential superfamily: members of this fold have similar functions but different ATP-binding sites

1. Ribokinase-like (2)

    has extra strand located between strands 2 and 3

    1. Ribokinase-like (3)
    2. Hydroxyethylthiazole kinase (thz kinase) (1)

TABLE 1 (continued)

2. MurD-like peptide ligases, catalytic domain (2)
has extra strand located between strands 1 and 2
1. MurD/MurF (2)
2. Folylpolyglutamate synthetase (1)

73. Carbamate kinase-like (1)
3 layers: a/b/a; mixed (mainly parallel) beta-sheet of 8 strands, order 34215786; strand 8 is antiparallel to the rest
1. Carbamate kinase-like (1)
topologically similar to the N-terminal domain of phosphoglycerate kinase
1. Carbamate kinase-like (2)

74. Class II aldolase (1)
3 layers: a/b/a; mixed (mostly antiparallel) beta-sheet of 9 strands, order 432159876; left-handed crossover between strands 4 and 5
1. Class II aldolase (1)
1. Class II aldolase (1)
metal (zinc)-ion dependent

75. Cytosolic phospholipase A2 catalytic domain (1)
3 layers: a/b/a; mixed beta-sheet of 9 strands, order 654321789; strands 4, 6 and 8 are antiparallel to the rest
1. Cytosolic phospholipase A2 catalytic domain (1)
1. Cytosolic phospholipase A2 catalytic domain (1)

76. Phosphatase/sulphatase (1)
3 layers: a/b/a; mixed beta-sheet of 10 strands, order 564371892A, (A=10) strand 9 is antiparallel to the rest
1. Phosphatase/sulphatase (2)
1. Alkaline phosphatase (1)
2. Arylsulfatase (2)

77. Isocitrate & isopropylmalate dehydrogenases (1)
consists of two intertwined (sub)domains related by pseudodyad; duplication

TABLE 1 (continued)

3 layers: a/b/a; single mixed beta-sheet of 10 strands, order 213A945867 (A=10); strands from 5 to 9 are antiparallel to the rest

1. Isocitrate & isopropylmalate dehydrogenases (1)
    1. Isocitrate & isopropylmalate dehydrogenases (7)

78. ATC-like (2)

consists of two similar domains related by pseudodyad; duplication

core: 3 layers, a/b/a, parallel beta-sheet of 4 strands, order 2134

1. Aspartate/ornithine carbamoyltransferase (1)
    1. Aspartate/ornithine carbamoyltransferase (6)
2. Glutamate racemase (1)
    1. Glutamate racemase (1)

    C-terminal extension is added to the N-terminal domain

79. Tryptophan synthase beta subunit-like PLP-dependent enzymes (1)

consists of two similar domains related by pseudodyad; duplication

core: 3 layers, a/b/a; parallel beta-sheet of 4 strands, order 3214

1. Tryptophan synthase beta subunit-like PLP-dependent enzymes (1)
    1. Tryptophan synthase beta subunit-like PLP-dependent enzymes (4)

80. SIS domain (1)

consists of two similar domains related by pseudodyad; duplication

3 layers: a/b/a; parallel beta-sheet of 5 strands, order 21345

1. SIS domain (2)
    1. "Isomerase domain" of glucosamine 6-phosphate synthase (GLMS) (1)
    2. Phosphoglucose isomerase, PGI (2)

    permutation of the superfamily fold

81. Formate dehydrogenase/DMSO reductase, domains 1-3 (1)

contains of two similar intertwined domains related by pseudodyad; duplication

core: 3 layers: a/b/a; parallel beta-sheet of 5 strands, order 32451

TABLE 1 (continued)

1. Formate dehydrogenase/DMSO reductase, domains 1-3 (1)
molybdopterine enzyme
1. Formate dehydrogenase/DMSO reductase, domains 1-3 (6)
domain 1 (residues 1-55) binds Fe4S4 cluster in FDH but not DMSO reductase

82. Aldehyde reductase (dehydrogenase), ALDH (1)
consists of two similar domains with 3 layers (a/b/a) each; duplication core: parallel beta-sheet of 5 strands, order 32145
1. Aldehyde reductase (dehydrogenase), ALDH (1)
binds NAD differently from other NAD(P)-dependent oxidoreductases
1. Aldehyde reductase (dehydrogenase), ALDH (8)

83. Aconitase, first 3 domains (1)
consists of three similar domains with 3 layers (a/b/a) each; duplication core: parallel beta-sheet of 5 strands, order 32145
1. Aconitase, first 3 domains (1)
1. Aconitase, first 3 domains (2)
contains Fe(4)-S(4) cluster

84. Phosphoglucomutase, first 3 domains (1)
consists of three similar domains with 3 layers (a/b/a) each; duplication core: mixed beta-sheet of 4 strands, order 2134, strand 4 is antiparallel to the rest
1. Phosphoglucomutase, first 3 domains (1)
1. Phosphoglucomutase, first 3 domains (1)

85. L-fucose isomerase, N-terminal and second domains (1)
consists of two domains of similar topology, 3 layers (a/b/a) each
Domain 1 (1-173) has parallel beta-sheet of 5 strands, order 21345
Domain 2 (174-355) has parallel beta-sheet of 4 strands, order 2134
1. L-fucose isomerase, N-terminal and second domains (1)
1. L-fucose isomerase, N-terminal and second domains (1)

86. Phosphoglycerate kinase (1)

TABLE 1 (continued)

consists of two non-similar domains, 3 layers (a/b/a) each
Domain 1 has parallel beta-sheet of 6 strands, order 342156
Domain 2 has parallel beta-sheet of 6 strands, order 321456
1. Phosphoglycerate kinase (1)
    1. Phosphoglycerate kinase (4)
        Domain 2 binds ATP

87. UDP-Glycosyltransferase/glycogen phosphorylase (1)
consists of two non-similar domains with 3 layers (a/b/a) each
domain 1: parallel beta-sheet of 7 strands, order 3214567
domain 2: parallel beta-sheet of 6 strands, order 321456
1. UDP-Glycosyltransferase/glycogen phosphorylase (4)
    1. beta-Glucosyltransferase (DNA-modifying) (1)
    2. Peptidoglycan biosynthesys glycosyltransferase MurG (1)
    3. UDP-N-acetylglucosamine 2-epimerase (1)
    4. Oligosaccharide phosphorylase (4)

88. Glutaminase/Asparaginase (1)
consists of two non-similar alpha/beta domains, 3 layers (a/b/a) each
Domain 1 has mixed beta-sheet of 6 strands, order 213456, strand 6 is antiparallel to the rest; left-handed crossover connection between strands 4 and 5
Domain 2 has parallel beta-sheet of 4 strands, order 1234
1. Glutaminase/Asparaginase (1)
    1. Glutaminase/Asparaginase (5)

89. Phosphofructokinase (1)
consists of two non-similar domains, 3 layers (a/b/a) each
Domain 1 has mixed sheet of 7 strands, order 3214567; strands 3 & 7 are antiparallel to the rest
Domain 2 has parallel sheet of 4 strands, order 2314
1. Phosphofructokinase (1)
    1. Phosphofructokinase (2)

TABLE 1 (continued)

Domain 1 binds ATP

90. Cobalt precorrin-4 methyltransferase CbiF (1)

consists of two non-similar domains

Domain 1 has antiparallel sheet of 5 strands, order 32415

Domain 2 has mixed sheet of 5 strands, order 12534; strands 4 & 5 are antiparallel to the rest

1. Cobalt precorrin-4 methyltransferase CbiF (1)
    1. Cobalt precorrin-4 methyltransferase CbiF (1)

91. Phosphoenolpyruvate carboxykinase (ATP-oxaloacetate carboxy-liase) (1)

consists of two alpha/beta domains

duplication: the domains share an unusual fold of 2 helices and 6-stranded mixed sheet; beta(2)-alpha-beta(4)-alpha; order 312465, strands 1 and 5 are antiparallel to the rest

1. Phosphoenolpyruvate carboxykinase (ATP-oxaloacetate carboxy-liase) (1)

domain 2 contains the P-loop ATP-binding motif

1. Phosphoenolpyruvate carboxykinase (ATP-oxaloacetate carboxy-liase) (1)

92. Chelatase-like (2)

duplication: tandem repeat of two domains; 3 layers (a/b/a); parallel beta-sheet of 4 strands, order 2134

1. Chelatase (2)

interdomain linker is short; swapping of C-terminal helices between the two domains

    1. Ferrochelatase (1)
    2. Cobalt chelatase CbiK (1)

2. "Helical backbone" metal receptor (3)

contains a long alpha helical insertion in the interdomain linker

    1. Periplasmic ferric siderophore binding protein FhuD (1)
    2. TroA-like (2)

TABLE 1 (continued)

3. Nitrogenase iron-molybdenum protein (3)
contains three domains of this fold; "Helical backbone" holds domains 2 and 3

93. Periplasmic binding protein-like I (1)
consists of two similar intertwined domain with 3 layers (a/b/a) each: duplication
parallel beta-sheet of 6 strands, order 213456
1. Periplasmic binding protein-like I (1)
Similar in architecture to the superfamily II but partly differs in topology
1. L-arabinose binding protein-like (13)

94. Periplasmic binding protein-like II (1)
consists of two similar intertwined domain with 3 layers (a/b/a) each: duplication
mixed beta-sheet of 5 strands, order 21354; strand 5 is antiparallel to the rest
1. Periplasmic binding protein-like II (2)
Similar in architecture to the superfamily I but partly differs in topology
1. Phosphate binding protein-like (20)
2. Transferrin (8)
further duplication: composed of two two-domain lobes

95. Thiolase-like (1)
consists of two similar domains related by pseudodyad; duplication
3 layers: a/b/a; mixed beta-sheet of 5 strands, order 32451; strands 1 & 5 are antiparallel to the rest
1. Thiolase-like (2)
1. Thiolase-related (6)
2. Chalcone synthase (2)

96. Fe-only hydrogenase (1)
consist of two intertwined domains; contains partial duplication

TABLE 1 (continued)

1. Fe-only hydrogenase (1)

1. Fe-only hydrogenase (2)

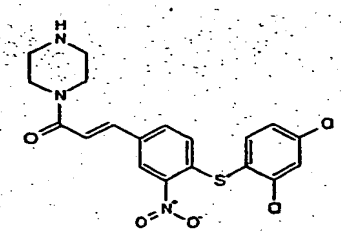
97. Cytidine deaminase (1)

consists of two very similar domains with 3 layers (a/b/a)each; duplication mixed beta-sheet of 4 strands, order 2134; strand 2 is antiparallel to the rest

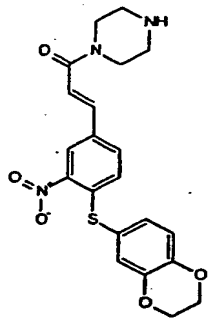
1. Cytidine deaminase (1)

1. Cytidine deaminase (1)

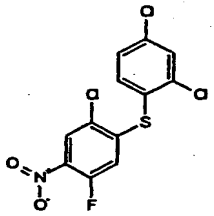
TABLE 2



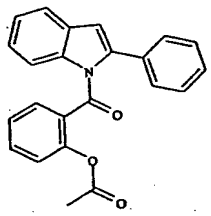
Cmpd A



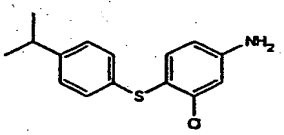
Cmpd B



Cmpd C

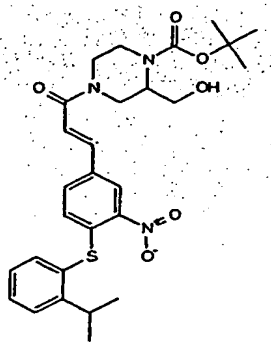


Cmpd D

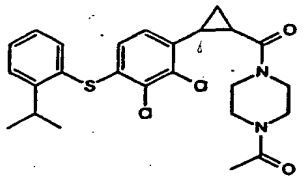


Cmpd E

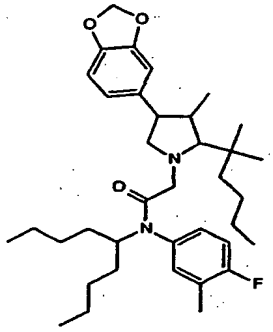
TABLE 2 (cont'd.)



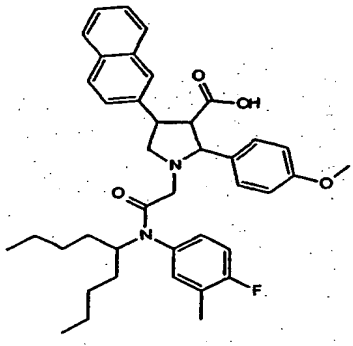
Cmpd F



Cmpd G

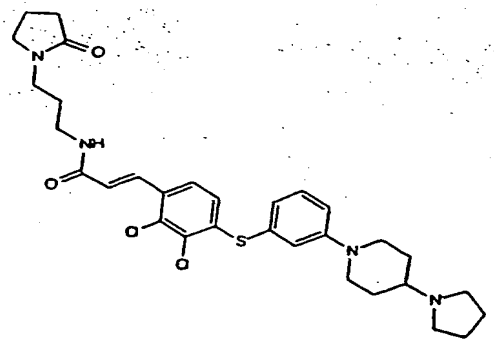


Cmpd H

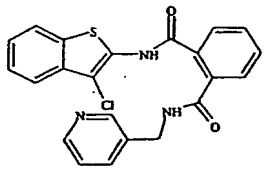


Cmpd I

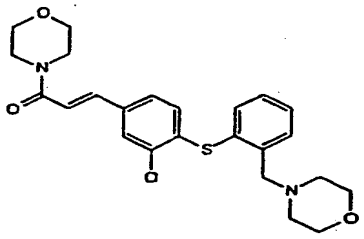
TABLE 2 (cont'd.)



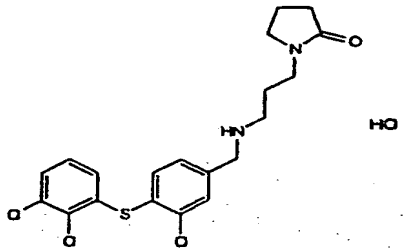
Cmpd J



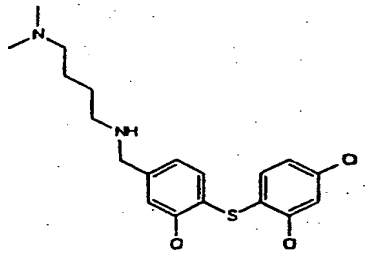
Cmpd K



Cmpd L

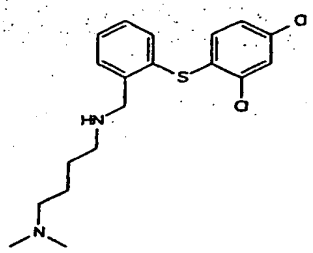


Cmpd M

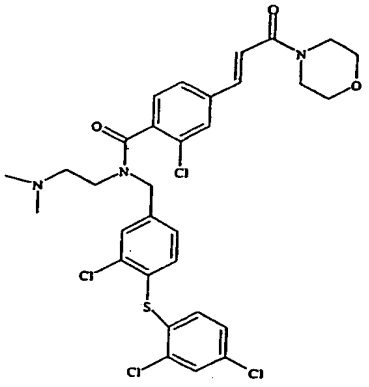


Cmpd N

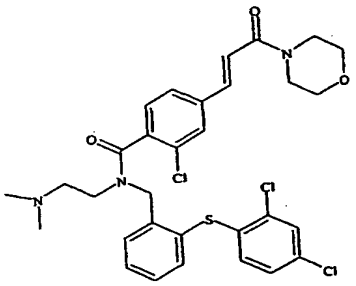
TABLE 2 (cont'd.)



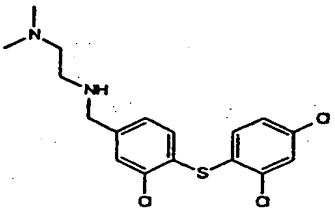
Cmpd O



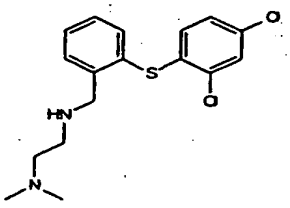
Cmpd P



Cmpd Q

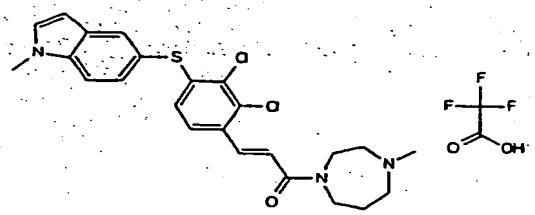


Cmpd R

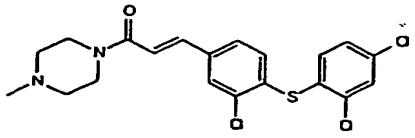


Cmpd S

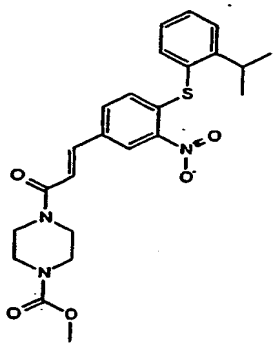
TABLE 2 (cont'd.)



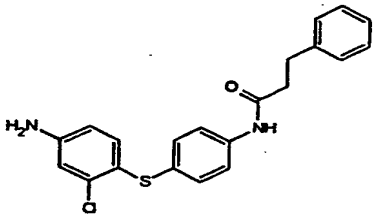
Cmpd T



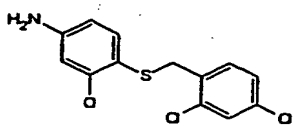
Cmpd U



Cmpd V

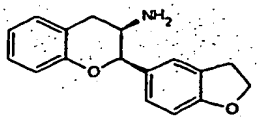


Cmpd W

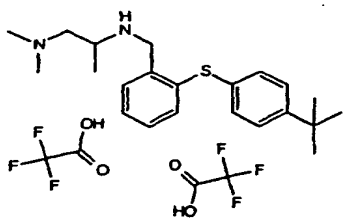


Cmpd X

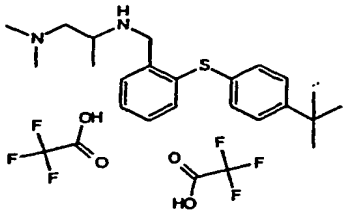
TABLE 2 (cont'd.)



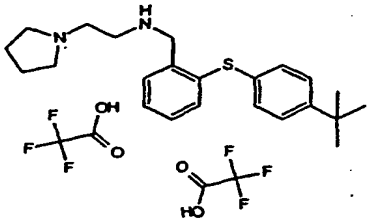
Cmpd Y



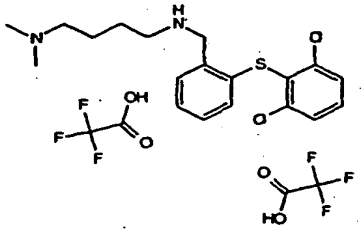
Cmpd Z



Cmpd AA

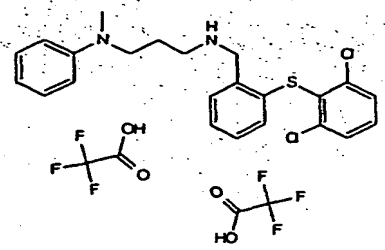


Cmpd AB

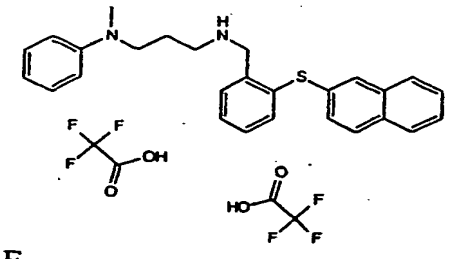


Cmpd AC

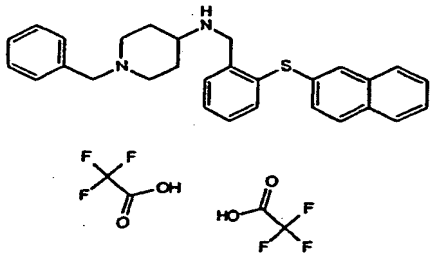
TABLE 2 (cont'd.)



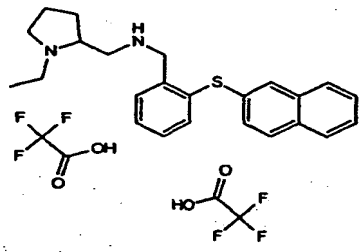
Cmpd AD



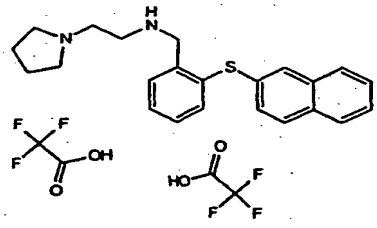
Cmpd AE



Cmpd AF

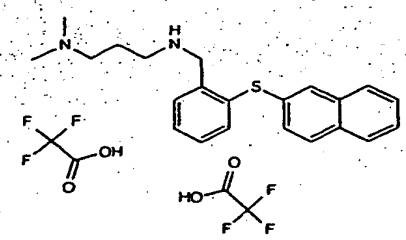


Cmpd AG

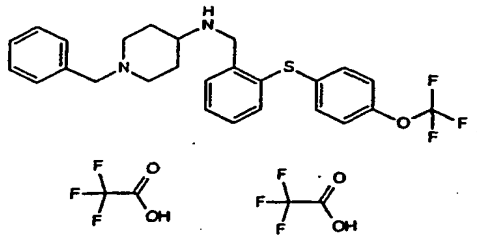


Cmpd AH

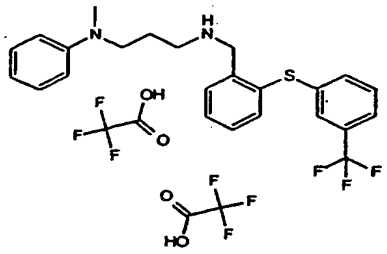
TABLE 2 (cont'd.)



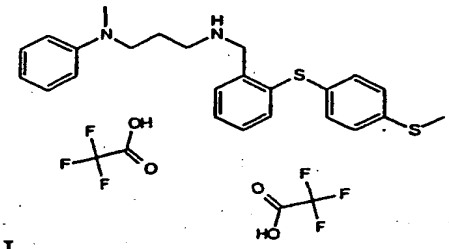
Cmpd AI



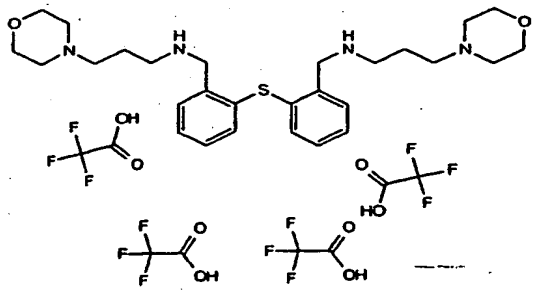
Cmpd AJ



Cmpd AK

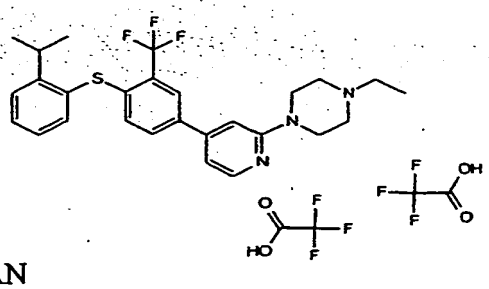


Cmpd AL



Cmpd AM

TABLE 2 (cont'd.)



Cmpd AN

TABLE 2 (cont'.d)

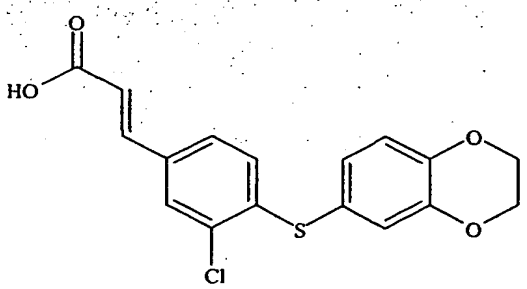
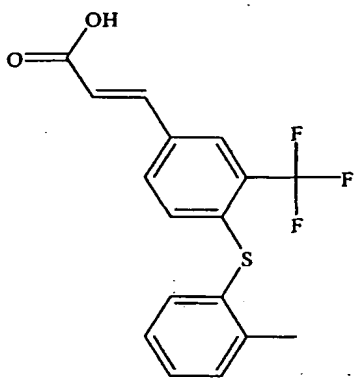
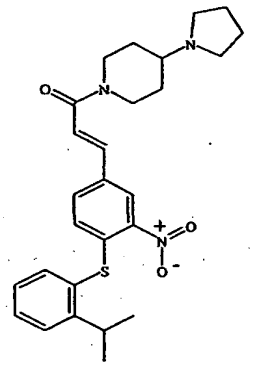
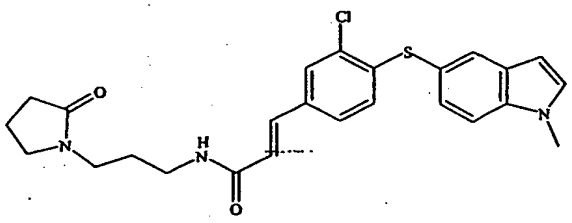
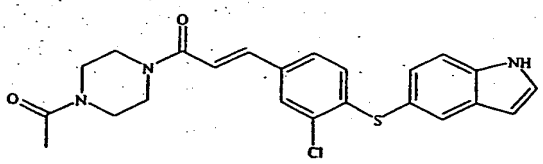
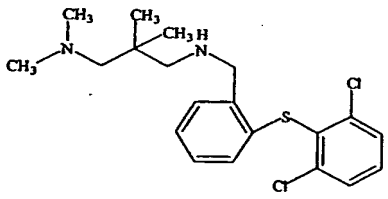
Cmpd AO 
Cmpd AP 
Cmpd AQ 
Cmpd AR 

TABLE 2 (cont'd.)

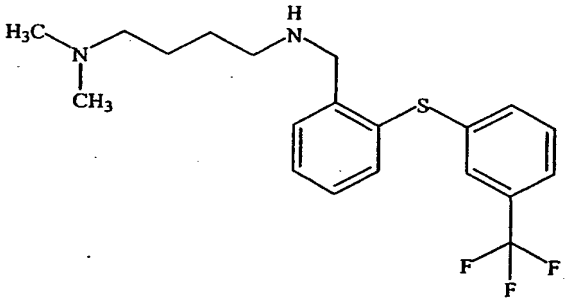
Cmpd AS



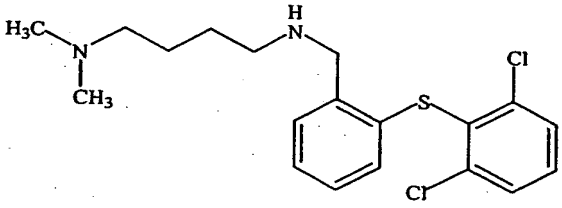
Cmpd AT



Cmpd AU



Cmpd AV



Cmpd AW

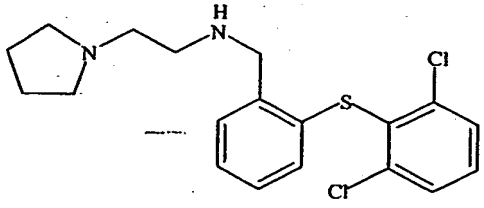


TABLE 2 (cont'd.)

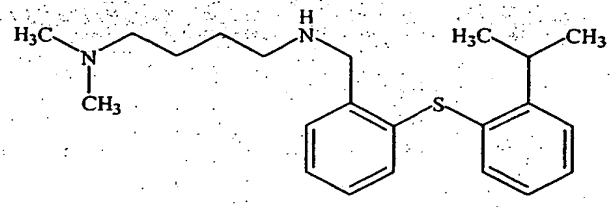
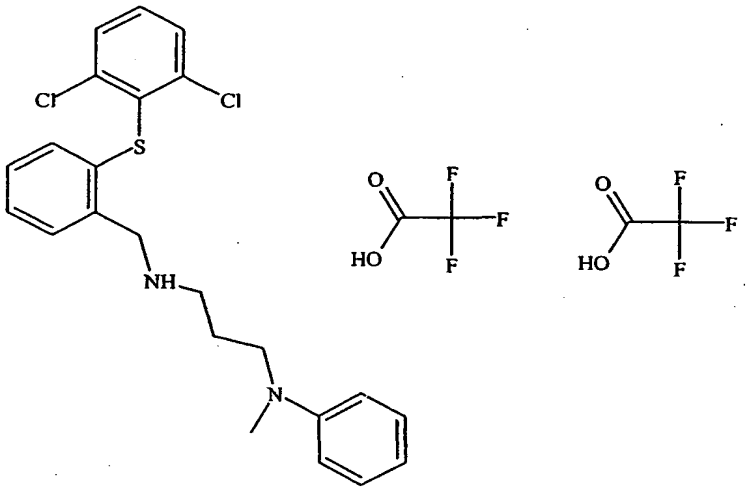
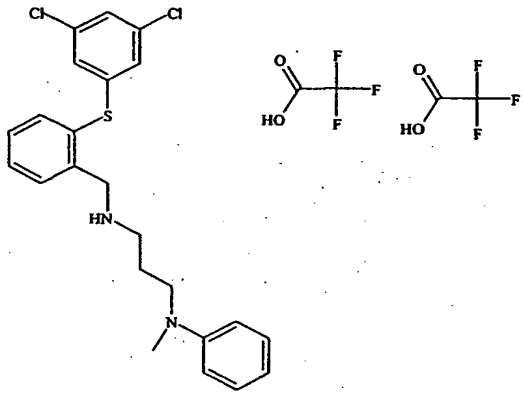
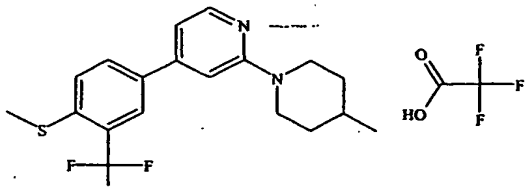
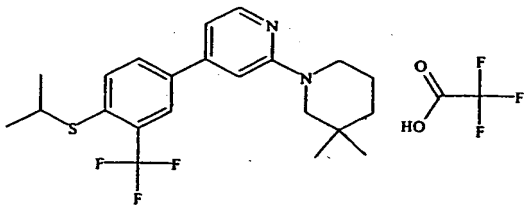
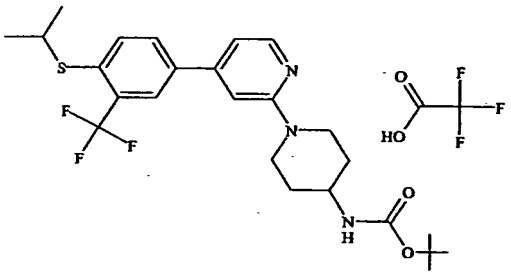
Cmpd AX 
Cmpd AY 
Cmpd AZ 
Cmpd AAA 

TABLE 2 (cont'd.)

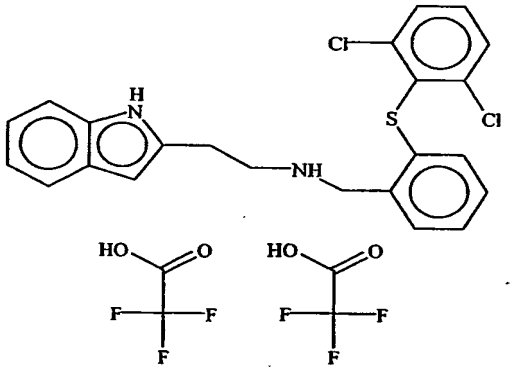
Cmpd AAB



Cmpd AAC



Cmpd AAD



Cmpd AAE

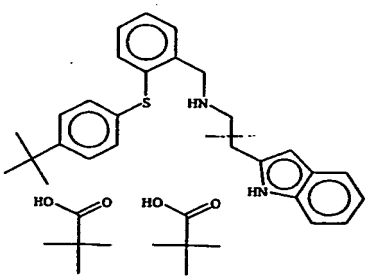
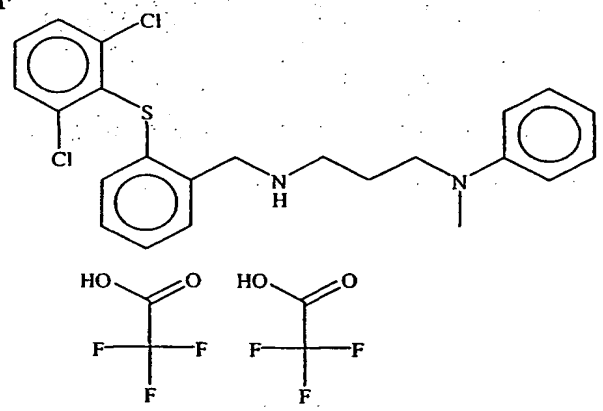
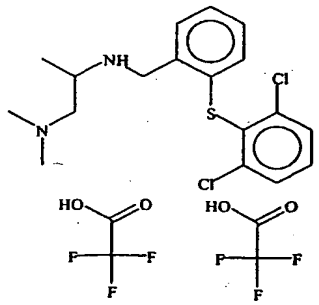


TABLE 2 (cont'd.)

Cmpd AAF 
Cmpd AAG 

The present invention is illustrated by the following examples.

Example 1

Identification of Alpha/Beta Proteins and Allosteric Regulatory Sites

The present invention also provides methods of identifying a molecule which is not LFA-1 or an I domain containing fragment thereof, said molecule comprising an α/β domain structure, said α/β structure comprising an allosteric regulatory site. When said molecule is contacted with an allosteric effector molecule, allosteric regulatory sites such as, for example, I domain allosteric sites, interact with said allosteric effector molecule to promote a conformation in a ligand binding domain of said α/β structure that modulates binding between the first molecule and a binding partner molecule thereof.

Allosteric regulatory sites can be identified, for example, by comparing candidate proteins to proteins having known allosteric regulatory sites. For example, α/β proteins having allosteric regulatory sites may be identified by using search tools, such as a NCBI vector alignment search tool (or "VAST" search), which are able to identify proteins similar to a predetermined three dimensional structure [Gibrat *et al.*, Curr. Opin. Struct. Biol. 6:377-385 (1996)], incorporated by reference herein in its entirety; and, Madej *et al.*, Proteins 23:356-369 (1995), incorporated by reference herein in its entirety]. With respect to these methods, LFA-1 can be used as a comparison or query protein because LFA-1 is known to include an I domain allosteric site. Similarly, other α/β proteins known to comprise an allosteric site can be used as a reference to identify other α/β proteins comprising an I domain allosteric site. In one embodiment, proteins with a VAST score of 7 or greater or a P value of 0.005 or less may be defined as being sufficiently related to the comparison protein to warrant further investigation.

Allosteric regulatory sites may also be identified by using an algorithm that predicts conformational ambivalence [Young *et al.*, Protein Science 8:1752-1764 (1999), incorporated by reference herein in its entirety; and, Kirshenbaum *et al.* Protein Science 8(9):1806-1815 (1999), incorporated by reference herein in its

entirety]. This algorithm, referred to as the Ambivalent Structure Predictor ("ASP"), predicts regions of three-dimensional conformational rearrangement from amino acid sequence information. The algorithm uses scaled probabilities from a secondary structural prediction algorithm, Profile Network Prediction Heidelberg ("PHD") [Rost, Meth. Enzymol. 266:525-539 (1996), incorporated by reference herein in its entirety], to identify structurally ambivalent sequence elements. Residues possessing a z score below -1.75 standard deviations of the mean residue ambivalence score in α/β domains are understood as being consistent with an allosteric regulatory site of the type useful according to the present invention.

For example, Table 3 shows that the integrin α/β domains and their close relatives possess a high VAST core of approximately 10 or greater and a P value of approximately 0.0009 or less relative to two representatives LFA-1 and Mac-1. Further, Table 3 indicates that the position of structurally ambivalent sequence elements (SASE) is consistent with the known or predicted c-terminal rigid body motion for these domains. Accordingly, these and other closely related domains of this type are predicted to possess a typical IDAS. Moreover, as demonstrated by the calculations presented in Table 3, some Ras superfamily members such as RhoA and enzymes such as ENR are also predicted to possess a typical IDAS.

Additionally, some non-integrin α/β domains that are more distantly related, as demonstrated by VAST analysis, possess a SASE at a site that appears to be distinct from the typical integrin IDAS. These α/β domains may possess an IDAS-like site also capable of being modulated with a small molecule such as a diaryl compound.

Many α/β domains share less than 35% amino acid identity. Therefore, a web-based simple modular architecture research tool, SMART, [see Schultz *et al.*, Nuc. Acids Res., 28:231-234 (2000), incorporated by reference herein in its entirety; Copley *et al.*, Curr. Opin. Struct. Biol. 9:408-415 (1999), incorporated by reference herein in its entirety; Ponting *et al.*, Nuc. Acids Res. 27:229-232 (1999), incorporated by reference herein in its entirety; and, Schultz *et al.*, PNAS USA 95:5857-5864 (1998), incorporated by reference herein in its entirety] that compares query sequences with its database of domain sequences has been used to identify additional divergent

family members. SMART utilizes multiple sequence alignments of representative family members. These alignments are optimized manually, and following the generation of a hidden Markov model, can be used to search sequence databases. Significantly similar sequences are added to the alignment, thereby refining the model which is used for subsequent searches. Accordingly, the SMART database may be used as a source of identifying additional α/β domains of interest to analyze for the presence of an allosteric regulatory site.

TABLE 3

αβ domain	VAST Structure Neighbor				ASP SASE* position (Residues from C-Termini)
	LFA-1		Mac-1		
	Score	P value	Score	P value	
α_L(LFA-1,1Z00)	–	–	14.7	10e-11.7	27
α_M(Mac-1, 1IDN)	13.2	10e-4.8	– .	--	28
α_1(1QC5)	13.8	10e-11.6	17.6	10e-15.9	23
α_2(1DZ1A)	12.5	10e-9.0	17.2	10e-15.3	16
ENR(IDFIA)	12.2	0.0009	10.8	0.0001	11
$G_{\alpha i}$(1GFI)			12.4	0.0016	70‡
Rac1(1MH1)			11.6	0.029	†
RhoA(1DPFA)			12.1	0.0045	23
cdc42(1AM4D)			11.6	0.253	20
H-Ras(1Q21)	10.4	0.0406	12.2	0.0027	†
Sir2(1ICIA)			8.0	0.0088	56‡
ftsZ(1FSZ)	11.7	0.0277	14.4	0.0048	92‡
HPPK(1DY3A)					37‡
Era (1EGA)	9.8	0.0474	13.3	0.001	81‡

*SASE: Structurally ambivalent sequence element.
† C-Terminal SASE not detected by ASP default settings.
‡ Second site of SASE may represent IDAS-like site.

Example 2

CD11b I Domain Mutants

A. Generation of Mutations in the CD11b I Domain

In view of previous results [Huth, *et al.*, Proc. Natl. Acad. Sci. (USA) 97:5231-5236 (2000)] using CD11a variants with mutations in the I domain, mutations were introduced in CD11b in an attempt to identify CD11b variants with increased affinity for binding partners ICAM-1and iC3b.

Six mutations were generated using a QuikChange Site-Directed Mutagenesis Kit (Stratagene). These mutants included single changes of Asp^{156} (D156A), Val^{254} (V254A), Gln^{327} (Q327A), Ile^{332} (I332A), Phe^{333} (F333A) and Glu^{336} (E336A) to Ala. Briefly, two mutagenic oligonucleotides (one to the sense strand and one to the antisense strand) were synthesized which were used in PCR with full-length CD11b as template. The PCR conditions for mutants D156A, V254A, Q327A, and I332A included 1 cycle at 95°C for 30 seconds followed by 16 cycles of 95°C for 30 seconds, 50°C for 1 minute and 60°C for 18 minutes. PCR conditions for mutants F333A and E336A were the same except that the final elongation step was carried out at 68°C for 20 minutes in the 16 cycles. After the PCR was complete, the methylated, non-mutated template DNA was digested with *Dpn*I at 37°C for 1 hour and the mutagenized CD11b DNA was used to transform Supercompetent XL1 Blue Cells (Stratagene) according to the manufacturer's suggested protocol. Carbomycin resistant colonies were picked and grown in liquid culture, after which plasmid DNA was isolated and the insert was sequenced. From clones having full-length mutants, a 1.3 kb *SacI*/*Eco*RV fragment containing the 5' portion of the gene was subcloned back into the parental vector. The inserts from these subclones were sequenced to verify the integrity of the junctions and the presence of the mutation.

D156A (sense) SEQ ID NO: 1
CATTGCCTTCTTGATTGCGGGCTCTGGTAGCATC

V254A (sense) SEQ ID NO: 2
GCCTTTAAGATCCTAGCGGTCATCACGGATGGAG

Q327A (sense) SEQ ID NO: 3
GAAGACCATTCAGAACGCGCTTCGGGAGAAGATC

I332A (sense) SEQ ID NO: 4
CAGCTTCGGGAGAAGGCGTTTGCGATCGAGGG

F333A (sense) SEQ ID NO: 5
CTTCGGGAGAAGATCGCGGCGATCGAGGGTAC

E336A (sense) SEQ ID NO: 6
GAAGATCTTTGCGATCGCGGGTACTCAGACAGG

B. <u>COS-7 Transfections</u>

COS cells were co-transfected with CD18/pDC1 and either wild-type CD11b or a mutant form of CD11b. Transfections were performed essentially as

previously described [Huth, *et al.*, Proc. Natl. Acad. Sci. (USA) 97:5231-5236 (2000)].

C. FACS Analysis

FACS analysis was carried out as previously described [Huth, *et al.*, Proc. Natl. Acad. Sci. (USA) 97:5231-5236 (2000)] except that the anti-CD11b monoclonal antibody TMG6-5 [Diamond, *et al.*, J. Cell Biology 120:1031-1043 (1993)] was used to confirm CD11b expression.

D. Adhesion Assay with COS Transfected Cells and Immobilized ICAM-1 or iC3b

Adhesion assays were performed in 96-well Easy Wash plates (Corning Glass, Corning, NY) using a modified procedure [Sadhu, *et al.*, *Cell Adhes. Commun.* 2:429-440 (1994)]. Each well was coated overnight at 4°C with 50 µl of glycophorin (Calbiochem) (10 µg/ml), ICAM-1/Fc (5 µg/ml), iC3b (3 µg/ml) or with anti-CD18 monoclonal antibody (TS1/18, 5 µg/ml) and anti-CD11b monoclonal antibody (44AACB [ATCC], 5 µg/ml) in 50 mM bicarbonate buffer (pH 9.6), or buffer alone. Plates were washed twice with 200 µl/well D-PBS and blocked with 1% HSA (100 µl/well) in D-PBS for 1 hr at room temperature. Wells were rinsed once with 100 µl of adhesion buffer (containing RPMI and 5.0% inactivated FBS) and 100 µl adhesion buffer was added to each well. Another 100 µl of adhesion buffer, with or without control antibody (IgG(5a)7.2, 60 µg/ml), blocking antibody (44AACB, 60 µg/ml) or activating antibody 240Q [Huth, *et al.*, Proc. Natl. Acad. Sci. (USA) 97:5231-5236 (2000)] at 60 µg/ml was added to each well, after which COS-7 transfectants (100 µl of 0.75 x 10^6 cells/ml) in adhesion buffer were added to each well. The plates were incubated at 37°C for 30 minutes for ICAM-1 binding or 15 minutes for iC3b binding. Adherent cells were fixed by the addition of 50 µl/well 14% glutaraldehyde in D-PBS and incubation continued at room temperature for 1.5 hr. The plates were washed with dH_2O, stained with 100 µl/well 0.5% crystal violet in 10% ethanol for 5 minutes at room temperature, and washed in several changes of dH_2O. After washing, 70% ethanol was added and adherent cells were quantitated by

determining absorbence at 570 nm and 410 nm using a SPECTRmax 250 microplate spectrophotometer system (Molecular Devices, Sunnyvale, CA). Percentage of cell binding was determined using the formula below.

$$\% \text{ of cell binding} = \frac{A570 - A410(\text{binding to ICAM-1 or iC3b})}{A570\text{-}A410 \text{ (binding to CD18+CD11b monoclonal antibodies)}} \times 100$$

Results indicated that wild type CD11b binding to ICAM-1 and iC3b was 3.1% and 26.4%, respectively. Mutants V254A, Q327A, and I332A each demonstrated significantly higher binding to ICAM-1 (114.7%, 105,1%, and 123.1% of wildtype levels, respectively) and iC3b (147.1%, 140.5%, and 205.2%, respectively), while mutants F332A and E336A showed significantly lower binding to both ICAM-1 (1.1% and 0.7%, respectively) and iC3b (4.9% and 4.3%, respectively). Mutants which demonstrate higher levels of ICAM-1 binding are therefore useful for identifying compounds that inhibit CD18/CD11b (Mac-1) binding to ICAM-1 in providing a higher signal-to-noise ratio as a result of the increased level of ICAM-1 binding.

Example 3
Identification of CD11b Agonists

Previous work has demonstrated that various diaryl compounds can inhibit LFA-1 binding to ICAM-1. In view of this observation and the results in Example 1 above, experiments were designed to determine if diaryl compounds can affect CD11b binding to natural binding partners, presumably through interaction with an allosteric regulatory region of CD11b.

A. Adhesion Assay of HL60 Expressing α_M to Immobilized ICAM-1

In order to assess the ability of the test compounds to modulate CD11b (α_M) binding, adhesion assays were performed using HL60 cells and immobilized ICAM-1.

Assays were performed in the presence of blocking anti-CD18 monoclonal antibody (TS1/22, 10 µg/ml) with 100 µl of HL60 cells (1 x 10^6 cells/ml)

in adhesion buffer were performed in 96-well Easy Wash plates (Corning Glass, Corning, NY) using the procedure described above except that each well was coated overnight at 4°C with (i) 50 µl ICAM-1/Fc (5 µg/ml), (ii) anti-CD18 monoclonal antibody (22F12C, 5 µg/ml) and anti-alpha 4 monoclonal (A4.1, 5 µg/ml) in 50 mM bicarbonate buffer (pH 9.6), or (iii) buffer alone. Percentage of cell binding was determined using the formula below.

$$\% \text{ Binding} = \frac{\text{A570 – A410(binding to ICAM-1)}}{\text{A570-A410 (binding to CD18+CD11a mAb)}} \times 100$$

Data was then normalized using the formula:

$$\% \text{ of DMSO binding} = \frac{\text{\% of cell binding, inhibitors}}{\text{\% of cell binding, DMSO}} \times 100$$

Approximately 30 compounds were identified for further study. IC50 values were determined in the HL-60 assay described above or in a neutrophil binding assays with fibrinogen described below (**Example 15**).

B. Adhesion Assay of JY/CD11b Cells to Immobilized iC3b

Briefly, each well of a 96-well plate was coated overnight at 4°C with 50 µl glycophorin (10 µg/ml), iC3b (5 µg/ml) or with anti-CD18 monoclonal antibody (22F12C, 5 µg/ml) and anti-CD11b monoclonal antibody (44AACB, 5 µg/ml) in bicarbonate buffer (pH 9.6). Plates were blocked with human serum albumin in D-PBS for one hr at room temperature. JY cells transfected with CD11b (JY/CD11b cells) (100 µl at 1 x 10^6 cells/ml) in adhesion buffer were added to each well and incubation was carried out at 37°C for 30 min. Plates were fixed and analyzed as described above in Example 1. Percentage of cells binding was determined using the equation below.

$$\% \text{ Binding} = \frac{\text{(A570 – A410(binding to iC3b)}}{\text{A570-A410 (binding to CD18 + CD11b mAbs)}} \times 100$$

Data was normalized using the formula:

$$\% \text{ of DMSO binding} = \frac{\% \text{ of cell binding, inhibitors}}{\% \text{ of cell binding, DMSO}} \times 100$$

IC50 values were determined for 45 compounds that demonstrated inhibition in the screen and six of these compounds showed IC50 of less than 10 μM. Twelve of the 45 compounds were subsequently used in binding assays using neutrophil adhesion to fibrinogen (described in Example 15).

This screen also identified 17 compounds with the ability to stimulate binding to iC3b. Re-titration of these 17 compounds revealed that Cmpd H, Cmpd I, and Cmpd C were capable of dose-dependent stimulation of CD11b/CD18 binding to iC3b at a level two times that observed with control DMSO treatment.

Example 4
Screening for Inhibitors of Complement Protein C2 and Factor B

Complement proteins C2 and Factor B have been shown to include A domain regions which are believed to regulate serine protease activity of the proteins and their respective convertases. The A domains in these proteins are also believed to serve as ligand binding sites and to include one or more regulatory domains. C2 binds complement protein C4b to form the C3 convertase and part of the C5 convertase in the classical complement pathway, and Factor B binds C3b to form the alternative complement pathway C3 convertase and part of the C5 convertase. Identification of modulators for C2 or Factor B binding would presumably provide a mechanism by which C3 and/or C5 convertase activity can be controlled.

A screen for inhibitors of the classical pathway complement protein C2 and alternative pathway complement protein Factor B includes primary screening using modifications of standard hemolytic CH50 and AH50 assays in a microtiter plate format as described below. [See also Current Protocols in Immunology, Chapter 13, Unit 13.1, John Wiley & Sons, Inc., (2000).] The CH50 assay is dependent on the activity of the classical pathway and C2, whereas the AH50 assay is dependent on the activity of the alternative pathway and Factor B.